T.R.

AKDENİZ UNIVERSITY

INSTITUTE OF EDUCATIONAL SCIENCES

DEPARTMENT OF EDUCATIONAL SCIENCES

MA THESIS

EXAMINING ONLINE TESTING PRACTICES IN FOREIGN LANGUAGE EDUCATION THROUGH THE PERCEPTIONS OF INSTRUCTORS AND STUDENTS: A CASE STUDY

KERİME ASMA

ENGLISH LANGUAGE TEACHING

MASTER'S PROGRAM

Antalya, 2021

AKDENIZ UNIVERSITY

INSTITUTE OF EDUCATIONAL SCIENCES

DEPARTMENT OF FOREIGN LANGUAGE EDUCATION

PROGRAM OF ENGLISH LANGUAGE TEACHING

EXAMINING ONLINE TESTING PRACTICES IN FOREIGN LANGUAGE EDUCATION THROUGH THE PERCEPTIONS OF INSTRUCTORS AND STUDENTS: A CASE STUDY

MASTER'S THESIS

Kerime ASMA

Supervisor

Prof. Dr. Binnur İLTER

Antalya, 2021

ATTESTATION

I honourably certify that the current study presented as my MA dissertation reports original work by me and was written within the scope of the ethics of research. I verify that the references contain the sources I utilized for my research through citation. I notify that I will bear all the ethical and legal consequences in condition that it is refuted.

DOĞRULUK BEYANI

Yüksek Lisans tezi olarak sunduğum bu çalışmayı, bilimsel ahlak ve geleneklere aykırı düşecek bir yol ve yardıma başvurmaksızın yazdığımı, yararlandığım eserlerin kaynakçalarda gösterilenlerden oluştuğunu ve bu eserleri her kullanışımda alıntı yaparak yararlandığımı belirtir; bunu onurumla doğrularım. Enstitü tarafından belli bir zamana bağlı olmaksızın, tezimle ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçlara katlanacağımı bildiririm.

16/08/2021

Kerime ASMA

ACKNOWLEDGEMENTS

This thesis has been shaped by the encouragement and cooperation of many people. By this way, I would like to state my thanks to those people.

I very much appreciate the cooperation and sincere responses of the participants of this study.

I would like to extend my heartfelt regards to my supervisor Prof. Dr. Binnur İlter for her invaluable guidance, inspiration, constructive feedback, assistance, motivation and patience throughout my study. Without her never-ending help and support, this study would not have been possible.

I would like to express my deep gratitude to the rest of the committee members; Assoc. Prof. Dr. Bengü ASKU ATAÇ and Assoc. Prof. Dr. Mustafa CANER for their proffessional guidance, significant contributions, useful suggestions and widening my research.

I would also like to thank my parents for their great support throughout my life.

Last but not the least, I owe special thanks to my dear life partner, my husband Burak, for his endless support, patience, respect and understanding.

Kerime ASMA

ABSTRACT

EXAMINING ONLINE TESTING PRACTICES IN FOREIGN LANGUAGE EDUCATION THROUGH THE PERCEPTIONS OF INSTRUCTORS AND STUDENTS: A CASE STUDY

Asma, Kerime

Master of Arts, Department of Foreign Language Education

Supervisor: Prof. Dr. Binnur İLTER

August, 2021, 107 pages

This study aims at examining online exam practices in foreign language education from the perspective of instructors and students and presenting a clear-cut picture of the current situation. The research was prepared in accordance with the case study design, one of the qualitative research methods. The study group of the research consists of 134 preparatory class students and 26 instructors in the preparatory school of a private university located in a province in Southern Turkey. The data of the study were collected using online data collection tools due to the Covid-19 pandemic, which started in December 2019 and is still in effect. For this purpose, separate opinion forms were prepared for both students and instructors. To analyse the collected data, quantitative descriptive and comparison analyses and content analysis technique were used. As a result of the analysis of the data, it was determined that the success levels of the students in face-to-face exams did not show a statistically significant difference according to gender, English level or computer skills, but there was an inverse relationship according to their age. While the stress levels of the students in face-to-face exams did not show a statistically significant difference according to age, English level or computer skills, they differed significantly in favour of male students according to their gender. Students' success in online exams did not show a statistically significant difference according to their gender and age; however, there was a significant difference in direct proportion to their English level and computer skills. While the stress levels of the students in online exams did not show a statistically significant difference according to their age or English level, they differed significantly according to their gender and computer skills. Accordingly, female students had greater stress levels than males; It was determined that students with low computer skills had a higher stress level than students with higher computer skills. As a result of the analysis of the qualitative data collected within the scope of the research, it was found that the students stated

the features such as creating a less stressful exam atmosphere, flexibility of the place and better concentration were the advantages of online exams; connection/technical problems and the stress associated with these problems, the exam environment being suitable for cheating, adaptation/concentration problems, finding a suitable place and difficulties in communication were the disadvantages of online exams. It was determined that the main problems faced by students in online exams were connection problems/technical problems and sound/noise problems. Students stated 18 factors that needed to be changed/improved regarding online exams, and the most mentioned factors were extending the exam duration and removing the rule to keep the microphone and camera on during the exam. In addition, 11 reasons were put forward as to why additional time should be given in online exams, the most frequently mentioned reason was the difference in connection and technical features. While most of the students stated that they showed their real performance in online exams, some students also stated that they performed better. The skill that students had the most difficulty in online exams was listening. They stated that if they had a choice, they would want online exams, and as a skill, they would most likely want speaking skills to continue to be measured online. As a result of the examination of the data of the instructors, it was seen that most of the instructors thought that the online exam practices in their institutions were practical by citing reasons such as preventing cheating, consisting of open-ended questions, being reliable and sufficient in current conditions. In addition, they stated that they had problems in four subjects: cheating issues, connection problems, difficulty in proctoring, and grading problems in online exams. The instructors also noted that online exams had advantages such as being less stressful, eco-friendly, easy to grade, flexible in terms of place, time-saving, preparing students for future professions and encouraging autonomous learning. They stated that the disadvantages of these exams were an environment conducive to cheating, inadequacy of social interaction, inability to measure students' skills, being time-consuming and tiring, stressful for both parties, constraint of using screen and technical problems. For alternative models/techniques for online testing, they made the following suggestions: project/portfolio evaluation could be conducted, students' computers could be blocked from opening different pages/browsers during the exam, plagiarism detection programs could be used, reliable electronic proctoring systems could be utilized, mirrors could be sent to students, different question types could be used, and these questions could be presented by shuffling.

ÖZET

ÖĞRETMEN VE ÖĞRENCİLERİN GÖRÜŞLERİYLE YABANCI DİL ÖĞRETİMDE ÇEVRİMİÇİ SINAV UYGULAMALARININ İNCELENMESİ: BİR DURUM ÇALIŞMASI

Yüksek Lisans, Yabancı Diller Eğitimi Anabilim Dalı

Danışman: Prof. Dr. Binnur İLTER

Ağustos, 2021, 107 sayfa

Bu araştırmanın amacı, yabancı dil öğretiminde çevrimiçi sınav uygulamalarının öğretmen ve öğrencilerin gözüyle incelemek ve mevcut durum hakkında betimsel bir resim ortaya koymaktadır. Araştırma, nitel araştırma yöntemlerinden durum çalışması desenine uygun olarak hazırlanmıştır. Araştırmanın çalışma grubu, Türkiye'nin güneyinde yer alan bir ilde bulunan özel bir üniversitenin hazırlık sınıfında eğitimlerine devam eden 134 hazırlık sınıfı öğrencisi ve 26 öğretim elemanından oluşmaktadır. Araştırmanın verileri, 2019 Aralık ayında başlayan ve hâlihazırda etkisini sürdüren Covid-19 pandemisi nedeniyle çevrimiçi veri toplama araçları kullanılarak toplanmıştır. Bu amaçla, hem öğrenciler hem de öğretim elemanları için ayrı ayrı görüş formları hazırlanmıştır. Araştırma kapsamında toplanan verilerin analizinde nicel betimsel ve karşılaştırma analizlerinden ve içerik analizi tekniğinden yararlanılmıştır. Verilerin analizi sonucunda, öğrencilerin yüz yüze sınavlardaki başarı düzeylerinin cinsiyete, İngilizce seviyesine veya bilgisayar becerilerine göre istatistiksel olarak anlamlı farklılaşma göstermediği, ancak yaşlarına göre ters oranlı bir ilişki olduğu tespit edilmiştir. Öğrencilerin yüz yüze sınavlardaki stres düzeyleri ise yaş, İngilizce seviyesi veya bilgisayar becerilerine göre istatistiksel olarak anlamlı bir farklılaşma göstermezken cinsiyetlerine göre erkek öğrencilerin lehine anlamlı bir şekilde farklılık göstermiştir. Öğrencilerin çevrimiçi sınavlardaki başarı durumları cinsiyetlerine ve yaşlarına göre istatistiksel olarak anlamlı bir farklılık göstermemiş, ancak İngilizce seviyelerine ve bilgisayar becerilerine göre doğru orantılı bir şekilde anlamlı farklılaşma göstermiştir. Öğrencilerin çevrimiçi sınavlardaki stres düzeyleri ise yaşlarına veya İngilizce seviyelerine göre istatiksel olarak anlamlı bir farklılaşma göstermezken cinsiyetlerine ve bilgisayar becerilerine göre anlamlı bir şekilde farklılaşmıştır. Buna göre, kadın öğrencilerin erkek öğrencilerden daha yüksek stres düzeyine sahip olduğu; bilgisayar becerisi düşük olan öğrencilerin ise daha üst düzey bilgisayar becerilerine sahip öğrencilerden daha yüksek stres seviyesine sahip olduğu tespit edilmiştir. Araştırma kapsamında toplanan nitel verilerin çözümlenmesi sonucunda öğrenciler, daha az stresli bir sınav atmosferi oluşturması, yer esnekliği ve daha iyi konsantrasyon sağlaması gibi özelliklerinin çevrimiçi sınavların avantajları olduğunu; bağlantı/teknik problemler ve bu problemlerle ilişkili olarak ortaya çıkan stres, kopya çekmeye

müsait sınav ortamı, adaptasyon/konsantrasyon problemleri, uygun yer bulma ve iletişimde zorluk gibi durumların ise çevrimiçi sınavların dezavantajları olduğunu ifade etmiştir. Öğrencilerin çevrimiçi sınavlarda karşılaştığı başlıca problemlerin bağlantı problemleri/teknik problemler ve ses/gürültü problemleri olduğu tespit edilmiştir. Öğrenciler çevrimiçi sınavlarla ilgili değiştirilmesi/geliştirilmesi gereken 18 faktör belirtmiş, bu faktörler arasında en çok belirtilenleri sınav süresinin uzatılması, mikrofon ve kameranın sınav esnasında açık tutulması kuralının kaldırılması olmuştur. Buna ek olarak, çevrimiçi sınavlarda neden ek süre verilmesi gerektiği konusunda 11 gerekçe ortaya koymuş, bu gerekçeler arasında en çok dile getirileni bağlantı ve teknik özelliklerdeki farklılıklar olmuştur. Öğrencilerin çoğu çevrimiçi sınavlarda gerçek performansını gösterdiğini ifade ederken bazı öğrenciler daha iyi performans gösterdiklerini de belirtmiştir. Öğrenciler çevrimiçi sınavlarda en çok zorlandığı beceri dinleme becerisi olduğunu; bir tercih hakları olsa çevrimiçi sınavları, beceri olarak ise en çok konuşma becerisinin çevrimiçi olarak ölçülmeye devam etmesini isteyeceklerini dile getirmiştir. Öğretim elemanlarının verilerinin incelenmesi sonucunda, öğretim elemanlarının çoğunun kurumlarındaki çevrimiçi sınav uygulamalarının kopyayı önleme, açık uçlu sorulardan oluşması, mevcut koşullarda yeterli ve güvenilir olması gibi gerekçeleri öne sürerek kullanışlı olduğunu düşündükleri görülmüştür. Ayrıca, öğretim elemanları çevrimiçi sınavlarda kopya olayları, bağlantı problemleri, sınav gözetiminin zorluğu ve puanlama problemleri olmak üzere dört konuda problemler yaşadıklarını ifade etmiştir. Öğretim elemanları, çevrimiçi sınavların daha az stresli olması, çevre dostu olması, kolay puanlanabilir olması, yer esnekliği ve zaman tasarrufu sağlaması, öğrencileri gelecek mesleklere hazırlaması ve otonom öğrenmeye teşvik etmesi gibi özelliklerinin avantajları olduğunu; kopyaya elverişli bir ortam yaratması, sosyal etkileşimin yetersizliği, öğrencilerin becerilerini ölçmede yetersizlik, zaman alıcı ve yorucu olması, her iki taraf için de stresli olması, ekran kullanma kısıtlılığı ve teknik sorunlar gibi unsurların ise bu sınavların dezavantajları olduğunu ifade etmiştir. Öğretim elemanları, çevrimiçi sınavlara yönelik alternatif olarak proje/portfolyo değerlendirmesinin yapılabileceğini, öğrencilerin sınav esnasında bilgisayarlarının kitlenerek farklı sayfaları/tarayıcıları açmasının engellenebileceğini, intihal tespit programlarından yararlanılabileceğini, güvenilir elektronik sınav gözetim sistemlerinden yararlanılabileceği, öğrencilere kullanmak üzere aynaların gönderilebileceği, farklı soru türlerinin kullanılabileceği ve bu soruların karıştırılarak öğrencilere sunulabileceği gibi önerilerde bulunmuştur.

Anahtar Kelimeler: Çevrimiçi Sınavlar, Ölçme ve Değerlendirme, Durum Çalışması, Üniversite Öğrencileri, Öğretim Elemanları

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

CHAPTER II

LITERATURE REVIEW AND RELEVANT STUDIES

CHAPTER III

METHODOLOGY

CHAPTER IV

FINDINGS

CHAPTER V

CONCLUSION, DISCUSSION AND SUGGESTIONS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFT: American Federation of Teachers

CB: Computer-based

CEFR: Common European Frame of Reference

DLAB: Defense Language Aptitude Battery

ELT: English Language Teaching

Hi-LAB: High-Level Language Aptitude Battery

MLAT: Modern Language Aptitude Test

MoNE: Ministry of National Education

NCAE: National Council on Assessment in Education

NEA: National Education Association

PB: Paper-based

TOEFL: Test of English as a Foreign Language

TSPS: Turkish Essay Scoring System

UZEMSS: Distance Education Centre Examination System

CHAPTER I

INTRODUCTION

1.1. Background of the study

As a result of globalization and technological advancements, the world has encountered many changes from past to present in numerous areas. In recent years, rapid changes and important developments have been observed in Turkey in basic areas such as technology which makes people's life easier, economy, politics and culture (Ercengiz, 2020). One of the areas affected by fast change and developments is educational institutions and the programs implemented in these institutions (Gelen & Beyazid, 2007). These shifts obviously have influenced the methods and approaches having been used in education.

Teachers have been obliged to keep pace with the modern technology and do essential modifications in their teaching. Therefore, explicit updates have been visible in the educational process recently. While the existing methods and approaches are being revised, brand-new techniques have come into existence. Contemporary practices in education aim at being appealing and motivating to the students, simultaneously preserving the educative side of the courses.

Since the unanticipated outbreak of COVID-19, the world has experienced the change to the digital world faster than ever. Like everything else, the continuity of educational progress has become feasible thanks to the advanced technology of today. Distance / Online education is an education model where educators and students are in different places and a communication path is established between them (Başkömürcü & Öztürk, 1996). Regardless of the place they are in, any student could continue his education with the co-existence of the internet and a technological device. There happen to be abundant debates on the effectiveness of the education provided during the pandemic circumstances, yet it is an undeniable fact that both parties – educators and learners- have been struggling and trying hard to succeed within the current state of the world.

In addition to all these updates in education, it was inevitable to experience up-to-date shifts in assessment and evaluation process of the educational practices. The developing world

conditions have also been decisive in the formation of new ideas in the field of assessment and evaluation and one of the most striking of these ideas is the online exam (Kip-Kayabaş, 2014). It is a well-known fact that assessment in education is an integral practice in education and serves critical data regarding the progress of the learners within the program offered at the institution. Assessment and evaluation, which is the last step of the stages in the educational programs; is the situations that determine how much of the behaviours desired to be seen in the student are acquired in line with the pre-determined goals in the education program, if they are not acquired what the reasons are, and what needs to be done to eliminate these setbacks, and the processes of making judgments based on these situations (Ornstein & Hunkins, 2009; Sönmez, 2015).

As in the whole world, the number of institutions providing distance education in Turkey has been increasing in parallel with the developing technology and increasing needs. Although many educational institutions were working on online programs to offer prior to the pandemic, it was an abrupt shift for testing practices. Initially, various methods and techniques were attempted to be utilized for online testing. Each educational establishment sought to discover a method accommodating best for their programs and throughout practice, they edited and adopted their own way to perform the exams online.

1.2. Purpose of the study

The purpose of this study is to determine the views of students and instructors on online exams and to present a clear-cut picture of the current situation. In line with this purpose, it was aimed at seeking answers to the following research questions in this study:

1) Is there a relationship between the students 'success and stress levels in face-to-face and online exams and their demographic characteristics (gender, age, English level, computer skills)?
2) What are the students' views on online testing practices in foreign language education?
3) What are the instructors' views on online testing practices in foreign language education?

1.3. Significance of the study

This research is important in terms of the characteristics of the study group that it examines. The online or distance teaching process, which became more important in our lives, especially with the Covid-19 pandemic, has started to be used more widely with the developing

and renewed education applications. Assessment and evaluation activities have a special place in these applications, which have become inevitable to be used in the future thanks to the advantages they offer. Although it already has some disadvantages, these disadvantages will be eliminated with systemic developments and updates, and it will be possible to carry out more effective and convenient assessment and evaluation activities. In this study, the current situation has been tried to be revealed by examining the views of students and instructors about online exams in the context of foreign language education in detail. As a result of carrying out similar studies in different fields, a holistic situation analysis for online exams will be obtained. In this way, these and similar studies will play a role as a guide in improving the system, eliminating its deficiencies and problems. When evaluated in this context, it can be said that this study is important for both foreign language education and other disciplines and fills the gap in the relevant literature.

1.4. Scope of the study

This study aims at determining the views of students and instructors on online exams and to present a clear-cut picture of the current situation. To attain this goal, a research design consisting of 134 preparatory school students and 26 instructors at a private university in Southern Turkey was created. In the study, the case study design, one of the qualitative research designs, was adopted. The research data were collected through online opinion forms in the 2020-2021 academic year. Separate forms were prepared for both students and instructors. The collected data were analysed through quantitative data analysis and content analysis technique. The findings obtained were presented in tables with examples of participant statements. Based on the findings, conclusions, discussions and suggestions were put forth.

1.5. Limitations of the study

The research data is limited to the data of the preparatory students studying in the preparatory class of a private university in the south of Turkey and the instructors who teach these students. In addition, data in the study were obtained using online data collection platforms due to the ongoing pandemic conditions. Finally, the case study design, one of the qualitative research designs, was adopted in the research, and qualitative data were emphasized in the collection of data.

1.6. Assumptions

The data of this research is based on student and instructor statements rather than official data. For this reason, it is assumed that the participants of the research answered the questions sincerely and intimately, and avoided comments and statements that would manipulate research findings.

1.7. Functional Definitions and Use of Concepts

Assessment / Measurement is observing a quality and showing the result of the observation with numbers or other symbols (Turgut, 1997). Determining the qualities to be measured, specifying the numbers and symbols suitable for this quality, matching the qualities and quantities, and making the assessment in accordance with the rules are respectively the stages of the assessment process (Baykul, 2015).

Evaluation, on the other hand, is the process of comparing the assessment result with a criterion and in this way reaching a decision about the feature determined by the assessment result (Özçelik, 2010). Evaluation, as understood from the definition, is based on three basic elements: the assessment result, the criterion, and the comparison and decision of these two elements. In this study, these terms are utilized interchangeably.

CHAPTER II

LITERATURE REVIEW AND RELEVANT STUDIES

2.1. Foreign Language Education

The main element of communication is language (Erdem & Eskimen, 2019) and it is a tool that has its own rules and provides communication among people within these rules. Aksan (2007) sees language as versatile and defines it as a highly developed system conveying thoughts, feelings and wishes to other people and language utilizes the units and rules which are common in a society in terms of sound and meaning. Besides, language is the main element that gives everything and its existence to humanity as a gift; the starting point of human adventure (Karaağaç, 2005). Language is what makes people privileged and powerful (Polat, 2006).

The language that the individual does not have the opportunity to use for communication with the environment he lives in and acquires for some purposes is called foreign language (Şahin, 2013). As it can be noticed, foreign language is acquired by the individual, not through environmental interaction, and learning a foreign language continues with a planned program in language education centers, preparatory programs of universities, schools, etc.

During the years of 1950s, after the Second World War, English became the most preferred foreign language in primary, secondary and higher education institutions in Turkey, as in the rest of the world (Bayyurt & Akcan, 2014). English is the most widely taught language in the foreign language teaching field in schools in Turkey (Çakır, 2018). This can be an indication of how important and international English has become as a foreign language. The major purpose of teaching English as a foreign language is to acquire and advance the students' communication skills, which consists of the ability to understand and express both in oral and written levels.

Teaching can be defined as the process of creating intellectual, emotional and physical change in the student in order to achieve a series of desired goals (Gronlund, 1985). The primary purpose of quality education is to train people with the power and equipment to change, and education is to train people with the knowledge and skills that can both adapt to the developments around them and solve the problems they may cause (Polat, 2006).

The characteristics of a student in the Foreign Language Curriculum are as follows (MoNE, 2017a):

- They constantly speak English in class.
- They actively participate in the lesson and constant interaction of the learners throughout communicative activities is expected.
- They use real-life English and use it effectively in different contexts consistently.
- They learn all of the four language skills integrally, in parallel with the acquisition of the mother tongue.
- Their in-class and out-of-class learning is encouraged so as to make them responsible individuals.

The characteristics of the teacher in the Foreign Language Curriculum are as follows (MoNE, 2017a):

- They act as role models for their students by constantly speaking English in class.
- They use different types of communication such as classwork, group work, pair work and individual work.
- They teach unfamiliar subjects in English by structuring them on the subjects students already know.
- They allow students to deduce meaning from context and/or from given clues.
- They show tolerance to students for mistakes and mispronunciations during speaking activities and use the correct form themselves, or they take notes to focus on the mistakes without giving the student's name after the activity for giving delayed feedback.
- They praise and give positive reinforcement to increase students' desire and motivation.
- Explain the rationale underlying language learning in general with specific language learning activities.
- They encourage, direct and guide students to learn English autonomously.

The critical threshold in foreign language education is intercultural interaction, in the meantime, the socio-cultural differences of individuals should be considered and an interactive language education model should be adopted (Yiğit, 2017). Some studies focus on alternative evaluation methods that consider individual differences (Çoruhlu, Nas, & Çepni, 2009); and result in a product and process-oriented evaluation (Yılmaz, 2018). Moreover, researches also

reveal that traditional product-oriented assessment methods for classifying learners are not curative (Başol, 2013; Çoruhlu et al., 2009; Kırık, 2008; Yılmaz, 2018).

2.2. Assessment and Evaluation in Language Learning/Teaching

The competency of a country in education shows both the development of that country and the place of its economy in the world. According to Tan (2014), when education is viewed with a system approach, the elements of the education system –like in every system- are input, process, output and evaluation. Therefore, the supplementary component of the educational process is assessment and evaluation activities. The answer to the question "How much has been learned?" at the end of the educational process is obtained through assessment and evaluation. As a result of this process, whether or not desired behaviours emerge or to what extent they are revealed, determining learning difficulties, specifying the effectiveness of education programs, methods and techniques, guiding students and evaluations for similar goals are all based on valid and reliable assessment results (Kan, 2017).

Assessment and evaluation processes, which have the quality of control over the education procedure, are included in every stage of the education. So as to be informed about the learning of the students, a teacher has to go through a three-stage process. The teacher first digitizes the answers of the students to the questions by utilizing the prepared test, then summarizes the results with tables or graphs to make the numbers more meaningful, and at the last stage makes a decision about the students (Berberoğlu, 2006; Çıkrıkçı-Demirtaşlı, 2014).

Assessment and evaluation consist of the basic concepts as assessment, evaluation and determining the situation. Assessment is the observation of whether a certain object or objects have a certain feature, and the degree of ownership, if any, and expressing the results of observation with symbols, especially with numbers. Evaluation is decision-making, and Turgut (1997) defined it as the job of making a decision by comparing the assessment results with a criterion.

In the educational process, various decisions are taken depending on the student's abilities, level of development and advancement potential. Determining these qualifications is a systematic assessment and evaluation process. The information obtained as a result of assessment and evaluation is utilized in making decisions about teaching, grading, guidance and counselling, curriculum, administration, and diagnosing students' development levels,

interests and abilities (Çepni et al., 2011) and decisions about selection and placement of students (Semerci, 2015).

With the help of assessment and evaluation processes, how successful the teacher is in decreasing the learning deficiencies of the less successful students, in gaining the teaching objectives, in guiding the students in the teaching process and in motivating the students are revealed.

In order to reach an accurate and useful value judgment, the assessment results must be valid and reliable, the evaluation must be based on a valid criterion, the evaluation procedures must be done without errors, and they must give a practical value judgment for the educational decisions concerned (Turgut & Baykul, 2015).

Concurrently, the most important data source of administrators, teachers and parents is the results of assessment which aids to determine the effectiveness of the knowledge and skills in implementing (Arıkan et al., 2014; Çepni et al., 2011). Five principles can be suggested to be selected and used for meaningful instructional assessments (Nitko & Brookhart, 2016):

1- Being clear about the learning objective to be evaluated,

2- Ensuring that the assessment techniques chosen match each learning objective,

3- Confirming that the needs of learners are served by the selected assessment techniques,

4- Assuring that various indicants of achievement are used for each learning objective as much as possible,

5- While interpreting the assessment results or helping students interpret them, it is necessary to insure that the limitations of these results are taken into account.

Representatives from the American Federation of Teachers (AFT), the National Education Association (NEA) and the National Council on Assessment in Education (NCAE) formed a committee and published a report in 1990 on the competencies that teachers should have in the field of assessment and evaluation. According to this report, there are seven assessment and evaluation competencies that teachers should have. Below are the competencies that teachers should possess (Grounlund & Waugh, 2009):

1- To be able to choose and develop suitable assessment and evaluation methods for education,

2- To be able to implement, score and interpret both externally sourced and self-produced assessment methods,

3- To be able to plan teaching, develop a curriculum, use the results of assessment on development of each student and school level,

4- To determine the status of students, being able to develop valid grading procedures,

5- To inform the students, parents, other educators and interested parties about the results of the evaluation,

6- To be able to understand unethical and illegal as well as inappropriate assessment methods and information.

In Turkey, in the study of the Ministry of National Education (MoNE) (2017) in which the general competencies of the teaching profession were determined, "Planning Education and Training", "Composing Learning Environments", "Administering the Teaching and Learning Process" and "Assessment and Evaluation" are stated as that teachers should possess these competencies within the professional skill competency area. The proficiency indicators within the assessment and evaluation field-competence mentioned in this report are:

1. Prepares and uses assessment and evaluation tools suitable for the field and developmental characteristics of students

2. Uses formative and summative methods in assessment and evaluation

3. Makes assessment and evaluation objectively and fairly

4. Gives accurate and constructive feedback to students and other stakeholders according to the assessment and evaluation results

5. Reorganizes teaching and learning processes according to assessment and evaluation results.

Foreign language teaching is an area where there are always great debates in assessment and evaluation and where these processes are mostly blamed by students, families and even teachers (Özmen & Balçıkanlı, 2012). This indicates that assessment and its tools have an important place in foreign language teaching as teaching language consists of four skills which are listening, speaking, writing and reading. The main purpose of the assessment is to ensure the improvement of the learning capacity of the student (Kesen, 2019) and to reveal effective and complete results about the success of the student. The validity and reliability of the prepared or selected tool are of great importance to ensure this. Moreover, assessment and evaluation play an active role in the progress of the foreign language education program in the targeted direction, determining its quality, seeing how much difference there is between the target point and the point reached, and recognizing the reasons for this gap (Şahin, 2018).

The development process of the students, their learning and language development levels are designated by the teachers via using various evaluation tools. A well-designed assessment system in foreign language education:

- Assists students discover what they know, understand and can do.
- Shows the students' level of development.
- Helps to plan the future learning process.
- Allows evaluating the standards that students are expected to reach in a certain period.
- Helps to describe how students can learn and do better.
- By sharing assessment results, provides information to the teacher, the student himself, and the parents about the student's learning process.
- Aids teachers and relevant people to make decisions about the implementation, monitoring and development process of the program.
- Supports the teacher while evaluating the adequacy of the methods and approaches used in the curriculum.
- Plays a critical role in identifying the areas that students have difficulty in comprehending, their weaknesses and knowledge gaps.
- Facilitates the teacher to design approaches and learning-teaching processes that will improve students' learning (MoNE. 2005).

Evaluation for foreign language teaching can be done both during and at the end of the course, or it can be provided for the whole class or just one person. Large scale events, formal or informal, are included in the assessment and evaluation. These can be both regular assignments and tests executed in certain periods (Abbott & Greenwood, 1985). Ceyhan (2007) states that the major aspect is to give the student the pleasure of success in exams. According to the constructivist approach, assessment should both contribute to the student's learning and enable the teacher to have an idea about the student's current thoughts and knowledge. Evaluation should not be a tool that pleases some students and worries others (Çakıcı, Gemici & Özsevgeç, 2008). Assessment and evaluation studies in foreign language education can be done in various ways. The vital part is to choose the type of assessment that suits the learners' needs most.

It is expected from foreign language teachers to have acquired an assessment and evaluation skill that can lead to healthy and reliable results in different areas of the language teaching process. As a feedback source, the assessment and evaluation phase should not be seen only as a tool to conclude the success level of students. Assessment and evaluation are resources to be used in determining both the effectiveness of the activities that teacher benefits from and

the efficiency of the educational process. In summary, assessment and evaluation guide the flow of the educational process. It concretizes the deficiencies of the process and also points out in which areas alterations are needed.

2.3. Types of Language Assessment

2.3.1. Formative Assessment

Throughout teaching, assessment and evaluation processes are carried out in order to constantly monitor the development of the students and to identify the learning deficiencies and problems. This type of process is known as formative assessment (Semerci, 2015). It is performed to identify and remove students' learning weaknesses, to enhance their learning. As Özçelik (2010) stated, it is a type of evaluation aiming at monitoring the students' progress and made at the end of each unit of the course to comprehend whether or not each student has learned what they need to master in each unit, and also what their deficiencies are. This assessment must be completed before each student moves on to the next unit or learning area. The purpose of diagnostic evaluation is to determine the causes of existing learning problems and to organize plans for their solutions (Linn & Gronlund, 1995; Şimşek, 2013). It provides continuous feedback to the program and in this type of assessment, a control system is established in order to take remedial measures (Demirel, 1999).

Evaluation occurs within a process and there is a continuous circulation in it. This type of assessment should be seen as a part of the teaching process. The main function of it is to identify learning difficulties and difficulties in each unit as teaching continues. The purpose of formative assessment is to guide learning and teaching (Harlen, 2005). It is a form of evaluation that guides the teacher with the feedback received from the student and helps him/her find the deficiencies in the student and draw a path accordingly.

Student success should not be evaluated for grading or other purposes depending on the results obtained. Increasing the efficiency of teaching and learning is the main purpose of this assessment.

2.3.2. Summative Assessment

The teaching process occasionally requires making the decisions at the end. In summative assessment, many decisions are made based on the grades given (Semerci, 2015). It is a type of assessment that evaluates the acquired behaviour, skills and characteristics of the

students and it is executed at the end of the program or course. The purpose of this evaluation is to determine the level that the students have reached in the process of the education, the degree of the students in reaching the goals of the educational programs and the power of the program to achieve this. (Yaşar, 2017). The purpose of the summative assessment is to grade the student and reach a decision about the student's proficiency level (Başol, 2013). This is mostly done by achievement tests. Thus, all the features that are desired to be acquired by the student are targeted to assess.

Generally, at the end of the teaching process (sometimes within the teaching process), decisions are made regarding the student, teacher and program, by examining whether the goals set by the program have been achieved or not. The data to be used in this type of evaluation are obtained by tests applied at the end of the course (general exam) embodying all the topics of the course or within the teaching period (midterm exam) including a few units (Tekin, 2000).

Another aim of this assessment type is to enhance and support student learning. As it is known; sensitive, accurate, appropriate and supportive assessment and evaluation is a requirement for learning (Tan, 2006). Therefore, teachers should be aware of the benefits of assessment and evaluation for themselves and their students.

As in all other studies, individual differences should be carefully and meticulously considered in assessment and evaluation activities. The aim is never to judge students, or to criticize them; the goal of carrying out this evaluation should be as an activity to support their academic, social or cultural development and guide them (MoNE. 2017).

2.3.3. Norm-Referenced Assessment

This is the evaluation made on the criteria determined according to the assessment results (Güler, 2018; Kaya & Semerci, 2017; Şahin, 2018). In norm-referenced assessment, each individual is evaluated according to the evaluation result of the class or group he is in, and thus his success in his class or group is determined (Nartgün, 2007). Norm-referenced assessment is commonly created to measure language skills, such as English proficiency, academic listening skills and reading comprehension (Brown, 2005). Norm-referenced exams can be used to see the differences among students and to form classes according to their level. The well-known TOEFL is a decent example of norm-referenced assessment.

Norm-referenced assessment includes the type of evaluation where each student's exam result/score is interpreted by comparing it with the scores of other students taking the exam. Norm-referenced assessment is interpreted through the concept of normal distribution. Student

scores are below or above the normal, and the success or failure is interpreted according to this. The "bell curve" evaluation method in which the passing grade is determined by taking the average grade of the class is an example of norm-referenced assessment (Atılgan, Kan & Doğan, 2017).

Followings are the vital features of the norm-referenced assessment:

- One student's acquisition is compared with all other students' percentage of their performance (in the exam)
- The student's skills and acquisition are measured
- It is aimed to ensure the continuity of the student's general success or abilities in the language
- It is quite long, contains many subtitles and covers a wide variety of content
- The student has little or almost no knowledge about the content of the questions that may be located in the exam (Brown, 2005).

2.3.4. Criterion-Referenced Assessment

The criterion-referenced assessment approach, which can be defined as the opposite of the norm-referenced assessment approach, involves the types of exams prepared with content and methods that reflect the objectives, content and practices of foreign language teaching programs, and in which student success is scored and interpreted with criteria determined according to the objectives. It is the evaluation made according to the criteria determined before the assessment process takes place (Güler, 2018; Şahin, 2018). In criterion-referenced assessment, each individual is assessed with his own success and the learners in a class or a group are evaluated independently from each other (Kaya & Semerci, 2017). As an illustration, a student who will be considered successful with a score of 60 or more, regardless of the scores of other students, is evaluated as norm-referenced.

Teaching objectives are determined depending on the course, the program, the education policy of the school and the state. The student's grade displays how much of the goals he acquired in terms of knowledge and skills without comparing with the grades of other students who took the exam (Davies, 1990). The purpose of criterion-referenced assessment is to obtain the student's level of comprehension of the subject taught in each lesson. While evaluating the knowledge or skills that the student has learned or acquired during the lessons, the focus is not on the distribution of scores, but on the knowledge and skills of each student (Bachman, 1990). In most cases, students know in advance what content questions and assignments they will

encounter in order to achieve each goal. Therefore, it is meaningful that the objectives related to the course content are also seen in the content of the questions (Bachman, 1990; McNamara, 2000).

Criterion-referenced assessment has many positive features. Firstly, certain skills and concepts are clarified. Also, the student's competencies and inadequacies in a particular field are explained. Secondly, one student's performance is not compared with another student's performance. Thirdly, the student's progress can be monitored constantly. Therefore, the performance of the student pre, while and post-teaching can be determined. Lastly, it allows the assessment of the level of achievement of the goals determined according to the needs of the student.

Below are the essential components of the criterion-referenced assessment:

✓ The extent to which a student has learned what was taught or the percentage of his success is determined.
✓ Defined goals specific to language learning are measured.
✓ It is aimed to see at what level each student comprehends what is taught.
✓ It has a short, well-defined structure that covers many subtests with similar content.
✓ The student knows exactly what the test content is (Brown, 2005).

2.3.5. Direct Assessment

Direct assessment is to directly measure or observe the feature to be measured with an assessment tool that is directly related to it (Güler, 2018; Kırık, 2008; Şahin, 2018). If the feature to be measured can be observed directly or measured with an instrument of the same type, it is a direct assessment. This assessment type is also called basic assessment (Turgut & Baykul, 2015). It is the expression of the observation that results with numbers and symbols by directly observing the values of the variables subject to the evaluation without any other variable being interfered with (Atılgan, 2012). To illustrate, an example of direct assessment is the measurement of the height by meters and finding 185 cm.

It is not essential to have an assessment tool in this assessment where the quality of the measured and the characteristics of the assessment tool are the same. Measuring the length of any object with a measuring tool of the same quality is direct assessment, and the evaluation of features such as gender, which can be directly observed without any mediator, is also direct assessment (Güler, 2018).

2.3.6. Indirect Assessment

In some cases, we cannot assess the variables that we want to evaluate by directly observing. In this case, we need to observe the variable or feature that we want to assess with the help of another variable or feature (Atılgan, 2012). This type of evaluation is called indirect assessment. The indirect assessment process is the evaluation of a feature that cannot be done without the help of another variable. To illustrate, a student's intelligence test with 118 points and a person's academic success are indirect assessments.

Indirect assessments are generally conducted in education (Semerci, 2015). Although it is not possible to evaluate academic success directly, it is possible to assess it with tests. For instance, with the intent of determining the degree of the students' success for a course, the answers given to the questions (directed to the students while assessing) are accepted as the indication of the success of the students in the course (Yaşar, 2017). According to Güler (2018), the assessment of cognitive and affective behaviors in social sciences is indirect assessment. Additionally, it is vital to bear in mind that some characteristics that we take as indicators of intelligence are what we actually evaluate, not the intelligence of the student.

2.3.7. Discrete-Point Assessment

In discrete-point assessment, it is intended to test each element at a time, in other words, item by item (Hudges, 2003), which focuses on the assessment of the students' knowledge from various grammar forms (Hidri, 2018). The idea behind the discrete-point testing is that language is perceived as dissectable; therefore, its components can be assessed separately (Resla, 1996). Thus, the goal is to evaluate the student's language skills by using various sub-tests.

2.3.8. Integrative Assessment

Oller (1979) claims that language learning is a set of unified skills, thus testing them one by one is not plausible. That is to utter that learners are to be assessed with their multiple linguistic competences simultaneously. Integrative assessment involves a combination of many different language elements in order to succeed in the test. Cloze tests, in which overall success of the student is to be assessed, can be a good illustration for integrative assessment. Likewise, by virtue of integrative assessment, a combination of different skills is assessed with the aim of evaluating the ability of the students to comprehend and internalize the information (Hidri, 2018). Hence, integrative assessment values all the skills equally. It is advantageous to utilize

this assessment type in that it has high validity and reliability and also enables grading the students objectively (Eckes & Grotjahn, 2006).

2.4. Principles of Language Assessment

2.4.1. Validity

In order to make a meaningful evaluation, it is necessary to refer to validity. Validity is a concept related to the characteristics of test scores and indicates the consistency of students' assessment results and their use. This term refers to the convenience of the assessment tool to the targeted feature to be assessed (Göçer, 2018). Stated in other words, a test contains validity in the event that it accurately measures what it is supposed to measure (Brown, 2004; Coombe, 2018; Rogier, 2014). Validity is a concept related to how accurately the test measures the characteristic of the individual (Büyüköztürk, 2016) and attributes to students' assessment results and the consistency of their use of those outcomes (Nitko & Brookhart, 2016). Baird et al. (2013) define it as the main conceptual tool used by assessment professionals to question evaluation policy and practices.

It is declared that the validity of the exam extends providing that the questions, supplied as an assessment tool, contain all the topics and the number is determined with a balanced distribution (Güngör, 2001). In order for the information to be collected for the development of the students to be as accurate and valid as possible, during the implementation of the exams, teachers should take into account the physical conditions of the classroom, the positive attitude and motivation of the students to answer the test, whether the day and time of the test is appropriate, and whether the students are affected by any external factors. (Berberoğlu, 2006).

There are some points that should be considered when questioning the validity of the assessment results. These are:

1. The concept of validity involves the ways of interpreting and using the results of the evaluation, not the assessment procedure per se.

2. The validity of an assessment tool is limited to the group and purpose from which it is obtained. Therefore, evaluation results have different degrees of validity for different situations and purposes.

3. Only after several types of validity evidence have been studied and put together should a decision be made about the validity of the use or interpretation of the evaluation results.

4. Validity is the judgment reached only after obtaining evidence from all relevant areas.

2.4.2. Reliability

Reliability is described as an indispensable characteristic of assessment. Reliability, in its shortest definition, means consistency in assessment. This term is a concept that expresses assessment results' degree of not containing random errors (Baykul, 2015; Tekin, 2000; Turgut, 1997). In another sense, Turgut & Baykul (2015) defines reliability as the degree of sensitivity of the assessments. Besides, Güler (2015) states that reliability can also be expressed with concepts such as consistency, stability and sensitivity.

Reliability, like validity, is a notion related to the scores or results obtained from assessment tools. In order for the decisions made based on the results that are obtained from the assessment tools to be accurate, the reliability of the assessment results gains importance. In other words, reliability shows the level of non-differentiation in student evaluations when students take the same test at different times, when the same task is scored by different teachers, and when different equivalent tests are administered at the same or different times (Yılmaz, 2011). Although the reliability value cannot be determined precisely, it is a value that can be estimated (Çıkrıkçı Demirtaşlı, 2014; Nitko & Brookhart, 2016).

Showing the same stability in every assessment indicates that the assessment tool has the quality of reliability (Ercan & Kan, 2004; Göçer, 2018; Özbek, 2017). Furthermore, reliability, which is expressed as the consistency of test results in assessment and evaluation, is persistent when the same test yields similar results in different implementations (Tafazoli, 2018).

As the number of questions in the assessment tool increases, the reliability parameter also increases (Şahin, 2018). The reliability of a test is measured with indices such as *Split-Half Method, Kuder-Richardson Confidence Parameter, Cronbach Alpha Confidence Parameter, Theta Confidence Parameter, Omega Confidence Parameter and Guttman Confidence Parameter* (Ercan & Kan, 2004). Reliability coefficient (a number between +2, -2) or standard errors (degree and amount of change between students) are reported when explaining reliability levels (Yılmaz, 2011).

Reliability requires a number of statistical processing and is determined by communication methods. Hence, sufficient evidence cannot be presented about the reliability of the evaluation results by making a logical analysis.

2.4.3. Practicality

Practicality (also called test practicality) refers to how practical, feasible and affordable a test is (Cohen & Swerdlik, 2013). On the other hand, Tekin (2000) defines practicality as the ease of development, reproduction, application and scoring of a test. From this viewpoint, practicality can be detached to the headings of affordability, scoreability and applicability.

Considering how long an exam takes, what kind of tools are used in the exam, how long the exam evaluation phase will take, how many people are required for the assessment, the cost of the exam (whether it is economical or not), the safe storage of the exam papers, etc. indicates whether it has practicality (Hughes, 2003; Clark & Lett, 1988; Brown, 1987).

The assessment tool that is more time- and cost-efficient in its development, implementation and evaluation phases has the feature of being practical (Güler, 2018). However, if a measure taken to increase the practicality of the tool will affect more important features such as reliability and validity, it would be more appropriate to abandon it (Karaca et al., 2014). Therefore, the practicality of an assessment tool depends on providing the right balance of reliability, validity, and practicality for its intended use (Tafazoli, 2018).

The practicality criterion only cannot be considered as sufficient for selecting an assessment tool/technique. For an exam, the validity and reliability qualifications are much more important than practicality. Especially, as the most important quality, validity is an indispensable element of an assessment and evaluation tool (Bachman & Palmer, 1996; Harmer, 2001).

2.4.4. Washback

It is the common opinion of the authors working in this field that the effect of foreign language exams on language teaching is very strong and that this effect, known as washback, is generally negative (Alderson & Banerjee, 2001; Brown, 2001; Brown, 1987; Hasselgreen, 2004; Hughes, 2003; McNamara, 2000). General assessments for selection and placement have a huge impact on education. The effects of exams, which are of great importance in the life of the student and his/her family in terms of their results, on the motivation of the student, classroom activities, school, education system and society in general, are obvious (Wall, 2005).

The results of the studies manifest that teachers do not focus on listening and speaking skills, which are not assessed in the exam, they tend more towards reading and writing activities, they use exam-oriented materials instead of textbooks, and they choose titles and text types of the exams (Wall, 2005). Teachers evaluate students according to the evaluation forms of the

previous exams and make the scoring in a similar way. In these studies, it was detected that the washback effects of the tests were reflected in the teaching content (Wall & Alderson, 1996). Turner (2001) defines the effectiveness of the reflection of the washback effects of exams on teaching content, techniques and activities as "learning shaped by exam techniques = exam-based teaching".

It is fairly important to look over the washback effects of tests on society in terms of the political and social dimensions. As an illustration, Davies (1990) claims that certain groups are disadvantaged because of their unequal background knowledge, and this is reinforced by exams. Shohamy (2001), on the other hand, emphasizes that education standards suitable for the middle class are established through examinations, that better job opportunities are provided in this way, low-income groups, minorities and immigrants are excluded through testing, and that this effect creates behavioral changes and disciplines these groups.

Major reasons for the negative effects of the washback can be test techniques and content, and the failure to provide reliability and validity criteria in the tests. Additionally, as stated by authors such as Davidson and Lynch (2008); Brown and Hudson (2002), the fact that the test content does not overlap with the curriculum goals causes the negative washback.

Discussions on how to eliminate the negative washback effects of exams have brought along discussions on alternative assessment and evaluation approaches and techniques, and methods have been sought for measures to be taken. Preparing exams with different content and techniques instead of monotype exams in assessment and evaluation practices may set an example for elimination. Therefore, the student is prevented from studying on the content and methods of the exam, and it is ensured that he works on what he needs to learn. The use of different techniques reflects real life and helps the student understand which techniques he expresses himself better. Thus, the student has a chance to discover his own learning and express it. Since the use of different types of assessment and evaluation techniques in the exams will naturally affect the teaching content and methods, it will enrich the teaching practices by saving them from monotypic.

2.4.5. Authenticity

Authenticity is the degree to which the characteristics of a particular language test task match those of a target language task (Bachman & Palmer, 1996). According to Brown (2001), as long as the language of the exam items is natural, items are context-based and tasks have

real-life situations, the assessment is authentic. In addition, the tasks should be correlated with daily life and appealing to the students.

Authenticity is correlated with construct validity, which is very significant in terms of validation, in that it supplies information about the degree of the generalization of the score interpretations (Bachman, 1990). Aside from this, students' perceptions and performance are influenced by authenticity. By taking these points into consideration, teachers' usage of authentic tasks for evaluation processes has great significance (East, 2008). Consequently, validity, reliability, practicality, washback and authenticity are the main principles of language assessment and evaluation; thus, it is required that teachers design the exams by keeping a balance among these principles.

2.5. Test types

2.5.1. Proficiency Tests

Hughes (2003) states that "Proficiency tests measure people's ability in a language, regardless of any training they may have had in that language." They are the types of exams that each institution, school or program prepares and evaluates in line with its own proficiency criteria. Exam content and question techniques are not determined according to the content and practices of any foreign language teaching program. Exam content and question formats are not standard and may vary according to each institution. Likewise, rating systems are non-standard.

Proficiency tests identify whether language is used creatively (Lowe, 1988). Broadly, proficiency exams, which are conducted to evaluate or determine the general knowledge and skills required or demanded when admitting students to educational institutions, are very general in nature and are not associated with the content and objectives of any language program (Hughes, 2003; Valette, 1977).

Proficiency decisions should be made on the basis of norm-referenced proficiency tests, as they require recognizing the student's overall proficiency level, which is derived in comparison with other students. Because norm-referenced evaluations have all the necessary qualifications to make such decisions (Bailey & Brown, 1999; Harmer, 2001).

Proficiency exams are carried out in order to decide according to which criteria the students joining the program will be placed or whether the students who want to transfer from a certain program to another program have achieved the necessary language proficiency (Brown, 2005; Brown, 1987; Herzog, 1988). Aside from the latter, at the request of centers

responsible for language education, proficiency exams are frequently performed in an attempt to compare schools or certain education systems.

2.5.2. Placement Tests

Placement tests are exams that foreign language teaching programs subject students with different knowledge and skill levels. In order to determine student needs in the context of qualification, the student's knowledge and skill level need to be clarified. Placement tests are held in order to make decisions about which knowledge/skill level the student is and to place students with similar knowledge/skill levels in the same classes.

Here, the main aim is not to assign grades to the student, as it is to ensure that the student is placed in the most proper program (or class) according to the grade he/she has received (Güler, 2018). In other words, there is no outcome as passed or failed.

Teachers benefit from this type of exam when they encounter students from different skill levels in the classroom. In that, it is important to create a class of students with similar knowledge and skills in order to determine the content and duration of the curriculum (Bachman, 1990; Bailey & Brown, 1999; Harmer, 2001; Hughes, 2003). Therefore, it is meaningful to compare placement tests with proficiency tests in order to clarify their qualifications. With the clarification of their functions and the obligations they undertake, at first glance, it is seen that the proficiency and placement exams are similar to each other since both tests assess general knowledge and skills.

As the proficiency tests are so general, they are created to evaluate a wide range of abilities/skills ranging from beginner to native language speaker level. Placement tests, however, are specifically linked to the established program and assess a very limited range of ability/skill areas and curriculum content. Thus, it effectively separates the students in the program according to the levels and determines the level of course content that the student can benefit most from. In this case, both proficiency and placement tests must be norm-referenced as it is more convenient to make conclusions according to the level determined by comparing the knowledge and skills of the student with the other students. Placement exams should be in accordance with the teaching goals and contents of the program administering the exam. However, proficiency tests are not exams created with this approach.

2.5.3. Diagnostic Test

This test type is implemented to define the pre-learning of students, which is to recognize their entry-level and to form a basis for the evaluation of their development. It is mostly executed to become familiarized with students upon their entry to the program. Ability, recognition, placement and exemption tests are the assessment tools used in this evaluation (Demirel, 1999). Conducting tests for students who would like to enrol in a foreign language course and placing the students in the appropriate levels according to the results of these tests can serve as examples for diagnostic tests.

The assessment and evaluation processes to be carried out before the learning process is undertaken to determine students' interests, readiness levels, abilities and to what extent students have acquired the knowledge, skills and attitudes that should be obtained at the previous learning (class, level or school) level and to place them the appropriate school, program or group. (Semerci, 2015). The aim is to improve the successful aspects of the student and to strengthen the weak and unsuccessful areas with supplementary support (Brown, 1987; Hughes, 2003). Diagnostic exams are criteria-referenced in terms of the function they perform. Besides, diagnostic tests can be applied throughout the education program, that is, during the learning and acquisition period of the language (Bachman, 1990; Davies, 1990; Harmer, 2001).

The purpose of such evaluations is to determine which branches, courses or teaching approaches will enable students to learn more efficiently by revealing personal characteristics and the development of students in different fields in detail (Özçelik, 2010).

At the end of this evaluation, it is feasible to determine the starting point of the teaching and to adjust the teaching according to the student's level. Compensatory teaching is planned if a deficiency in terms of entry behaviours is detected.

2.5.4. Achievement Tests

The type of exam utilized to determine the success and to see the student's needs and development levels is called achievement test. Achievement exam is the most appropriate assessment tool to see to what extent the student has successfully learned and acquired the targeted knowledge and skills. The most obvious feature of the achievement test is that it is administered at the end of the semester.

The aim of the achievement exams is to specify which student is ready to move on to a higher program or who can successfully complete the program and graduate (Brown, 1987; Hughes, 2003; Valette, 1977). Achievement tests serve as a good data provider to decide upon

how student success can be improved, what changes will be made while creating the curriculum, whether new recruitment is essential, what kind of activities, materials, tools, etc. are on-demand (Bachman, 1990; Bailey & Brown, 1999; Harmer, 2001; Harrison, 1983).

An achievement test is a vital source of information to make a decision on student success or to increase their success. It is also a requisite tool in terms of showing how much each student has learned throughout the program. Achievement tests should be in harmony with the objectives of the unit/school/class/program in which it is executed and therefore should be criterion-referenced. This test is administered at the end of the course/program to see how much of the student's teaching objectives have been achieved and successfully completed. According to the results obtained, achievement tests can also be used to make decisions about what changes should be made about the content, methods and nature of the activities utilized throughout the teaching process.

2.5.5. Aptitude Tests

In the literature, various definitions can be discovered regarding aptitude. To illustrate, Damon (2005) defines aptitude as character structure or personal characteristics that push people towards certain decisions and experiences. While Raths (2001) sees aptitude as closely related to skills and practices, Perkins and Tishman (1998), on the other hand, define the term aptitude as a preference to exhibit a behavior under certain conditions. Besides, the tendency to structure a frequent, conscious and voluntary behavior towards a goal is regarded as aptitude (Katz, 1993).

Regarding the scope of the current study, the term 'language aptitude' is required to be defined. Carroll (1981) explains the term as follows:

> "… foreign language aptitude is not exactly the same as what is commonly called 'intelligence,' not even 'verbal intelligence,' for foreign language aptitude measures do not share the same patterns of correlations with foreign language achievement as intelligence and academic ability measures have." (p. 86)

The potential capacity of the learners to acquire another language is aimed at discovering in language aptitude assessment. It is frequently a key aspect in terms of deciding whether it is deserving to spare time, endeavor and spend money on learning a second language (Doughty et al., 2010). The most widely-known and used language aptitude assessment tests are below:

- The Modern Language Aptitude Test (MLAT; Carroll & Sapon, 1959)

- Defense Language Aptitude Battery (DLAB; Lett et al., 2004; Peterson & Al-Haik, 1976)
- The High-Level Language Aptitude Battery (Hi-LAB; Doughty et al., 2010; Doughty, Campbell, Bunting, Bowles, & Haarmann, 2007)

2.6. Contemporary Testing Practices

2.6.1. Self-Assessment

The student's self-evaluation on a particular subject is called self-assessment. This approach helps students discover their talents (Ministry of National Education [MoNE], 2005). Self-assessment can be expressed as the student's judgment of his learning stages and success levels. The student takes responsibility for making judgments about his own learning and assesses himself. There is the possibility of student overestimating his competencies due to being biased. Hence, students should be assisted so that they can notice their competencies correctly and be unbiased, and students should gain this skill by constantly applying self-assessment (Alıcı, 2014; Tekindal, 2016).

2.6.2. Peer-Assessment

This is the process of evaluating the peers in a group (Kuyumcu & Erdoğan, 2007). In other words, peer assessment is defined as the process of evaluating another person or other people. The target in the practice of the assessment is to enable students to have a critical perspective and to ensure to give feedback to their friends by providing the necessary studies for them. (Buyer, 2014).

Since students are more likely to act subjectively in peer assessment, it is difficult to implement it objectively at the beginning. However, making this evaluation at intervals, explaining the significance of this activity to the students, providing guidance and informing the students about the evaluation criteria will be beneficial for objective assessment (Alıcı, 2014; Uysal, Öztürk & Döş, 2015).

2.6.3. Project

Projects are studies that students do individually or in groups in any field they wish. It is well-known that students interact with each other or with their environment within the scope of the project approach. Projects can be thought of as extended performance tasks that can be completed in a few weeks or even months. It is usually with the guidance of the teacher that the

students search in groups, make inferences for the purpose of interviews or obtain new information. Project topics aimed at students' daily lives and interests facilitate their studies (Yayla, 2012).

The biggest downside of the projects is the fact that they can cause rater errors. In order to minimize these errors, it is of great benefit to use a detailed rubric, as with other complementary assessment tools. In addition, the rubrics should be given to the student at the beginning of the process and the students should see the stages and the elements that will be evaluated (Turgut & Baykul, 2015).

2.6.4. Portfolio

CEFR (Common European Frame of Reference) which is a common framework aimed at determining language proficiency of the learners by referring to certain criteria founded in 2001, recommends portfolios as an assessment tool as they enable language learners to store their progress of learning experiences (CoE, 2001). Therefore, portfolios are a purposeful collection of studies or behaviors that reveal a meaningful portray of talents in any field (Alıcı, 2014). According to Aydoğdu and Kesercioğlu (2005), the portfolio is a collection of the activities of the students during the term, which are for a specific purpose, gathered in a file under the guidance of the teacher. Kılıç (2006), on the other hand, defined it as a file consisting of products that will display the student's knowledge, intended skills or achievements on a subject.

Portfolios constitute the development files that include the performance, development and studies that students put forward in order to serve a purpose in a certain process (Güler, 2018) and assist learners to be informed about their own improvement which would lead to self-assessment (CoE, 2018).

When the disadvantages of portfolios are considered, factors such as keeping them, not being able to know for sure whether the work in the file is done by the student himself, taking time to evaluate, and score reliability become prominent.

2.6.5. Performance Assessment

Performance evaluation is the practice made for students to learn actively, to assess the works that are carried out within a certain period of time and the products that emerge at the end of the process (Alıcı, 2014). It refers to the situations and assignments that will ensure that individual characteristics such as performance evaluation and learning types are taken into

account and that they are transformed into action (Ministry of National Education [MoNE], 2006). Broadly, we can define performance evaluation as activities and assignments that enable students to transfer the knowledge and skills they have learned into their own lives. Performance assignments are spread throughout the process and are not limited to a certain time.

2.6.6. Observation

Observation is a data collection tool that aims to examine human behavior and is used in the natural environment (Ekiz, 2009). This tool can be used to evaluate any observable performance of students in environments where complementary assessment and evaluation activities are applied in individual or group activities (Bahar, Nartgün, Durmuş & Bıçak, 2014). Teachers make use of the observation technique to obtain accurate and fast information about students and to unveil the reactions of students to the activities they do about learning (Uysal et al., 2015). It is a technique used to obtain accurate and primary information about students within the assessment and evaluation practices. In educational institutions, observations are made in a systematic and planned manner. A feature that distinguishes observation from other techniques is that it is an information-gathering technique that has been widely used from the past to the present. In the light of this information, it can be declared that the observation method can be applied in every lesson and environment.

2.6.7. Interview

Interview is a type of purposeful and planned data collection by communicating with individuals in various ways for certain purposes (Erkuş, 2006; Gümüş, 1977). In the interviews held in education, the answers given by the students are examined, and data are obtained about the extent to which that knowledge is acquired (Nartgün, 2014).

Interviews have a significant function in evaluating the level of understanding of students about their studies and course subjects (Uysal et al., 2015). It is possible to classify interview types as individual and group interviews. The questions to be asked in the interview can be prepared in advance or added and asked at that moment. The main purpose of using the interview technique is to understand students' experiences and how they make sense of these; not to test a hypothesis (Türnüklü, 2000).

2.6.8. Drama

Drama is to revive an experience, event, word, concept, sentence or thought by making use of theatrical activities such as role-playing (Özyürek, 2016). Drama is in a structure that uses the techniques and approaches of visual arts and learning-teaching principles in an integrated manner and needs materials such as text, scene and light, and requires participants to be active at every stage (Akar-Vural & Somers, 2012). The purpose of drama is to get children active in the learning process and to ensure the permanence of their gains by ensuring that they have a pleasant time throughout the process. In line with this purpose, the individual acquires an aesthetic point of view, critical thinking and expression, and the ability to look at events from multiple perspectives (İşyar, 2017).

In integrative learning theory, the individual constructs knowledge as a result of interaction with his environment. Therefore, in drama activities, instead of imitating knowledge; environments, where students will build knowledge themselves with their own experiences, are created for students (Çepni, 2005).

2.7. Distance/Online Testing

The assessment and evaluation process is an indispensable part of the learning process. Recently, the world has experienced a shift from face to face to online in testing practices as a result of the demand resulting from technological advancements. The number of studies on how technology can be used in assessment processes as well as in learning processes is increasing. Besides, the unexpected COVID-19 pandemic has obliged learners and educators to pursue this novel change in education. Moreover, technology-assisted assessment and evaluation can be more beneficial financially and also in terms of the workforce. When this process is supported with technology, extremely fast and reliable results are obtained. Online education includes different approaches, and therefore, online assessment and evaluation practices will be different from face-to-face education (Brown, 2004). For this purpose, Brown (2004) sought answers to the following questions in his study:

➤ Are the instructors sufficient to use online assessment methods?
➤ What are the assessment methods suitable for online learning processes?
➤ How should instructional design be structured in order to better evaluate success?

Continuous assessment and evaluation activities should be included in the online education process in order to control the learning speed of the students and to prevent any

disruptions or deficiencies in the learning quality to be obtained at the end of the evaluation process (Palloff & Pratt, 2009).

A well-designed online course and assessment need to be student-centered. Enabling the student to participate actively in learning activities such as discussion, group work and self-assessment in the online environment has an important place in the assessment phase of the educational process (Kanatlı, 2008). Some of the principles that teachers should consider when evaluating online can be summarized as follows:

➢ Planning student-centered assessments that include self-assessments
➢ To include joint studies where students can evaluate each other by making comments
➢ Preparing grading scales for assignments, projects and group work to aid assessment
➢ To utilize assessment techniques appropriate to the course objectives and content
➢ To include easy and clear-to-understand assessment methods and techniques that are enjoyable to work with online
➢ Communicating with the students on how the assessment should be and getting their ideas (Palloff & Pratt, 2003).

Quizzes and tests will of course be used for individual assessment (Morgan & O'Reilly, 1999) and in online environments, tests and quizzes are quite convenient. However, this is not a complete assessment. While evaluating the student in the virtual environment, the importance of using all assessment techniques should never be forgotten.

Angelo and Cross (1993) stated that when the evaluation process is studied together with the student, his comprehension of the course content and gaining self-assessment skills consolidate. Additionally, in student-centered teaching, determining a method by consulting how the student wants to be evaluated lays the groundwork for the student to show greater improvement (Bachman, 2000). Therefore, it would serve great benefits to get the opinions of the students about how the evaluations are carried out and then to add the data up to the evaluation draft.

The main advantages of online testing are stated below (JISC, 2007):

＋ Providing flexibility to students in terms of time
＋ Faster collection of data
＋ Rapid designation of results
＋ Cost reduction
＋ Interactive evaluation is possible

Motivation can be increased by using up-to-date technologies.

The limitations encountered in online testing are:

Exams require computer and internet access
Whether it is safe or not
There is a possibility of cheating by students or it is not possible to check whether the student himself took the exam
Lack of communication (Shuey, 2002).

Gaytan (2005) suggested the following techniques to create more efficient and effective assessments online:

Being in contact with the student regularly
Interaction through group work, cooperation and discussion should be kept at a dynamic level
Alternative assessments such as authentic assessments and e-portfolios should be employed.

2.8. Related Studies

When the related studies are examined, it was seen that the studies conducted with online exams were concentrated in certain categories. Among these categories, the most studied category was the studies in which online and paper-pencil exams were compared in terms of various variables. In these studies, exams were named in different ways. They were sometimes called online vs pen-paper, sometimes computer-assisted vs. paper-based, sometimes online vs. traditional, sometimes computer-assisted vs traditional. In the related literature, studies examining the opinions of students, pre-service teachers or teachers towards online exams were also noteworthy. Some studies examined the impact of online exams on subject areas such as students' exam performance, anxiety, and stress. Among the related studies, studies on proctoring were also determined. The last category related to this issue was the studies on the development and testing of the online exam/scoring system.

Table 2.1. *Related studies*

Studies on the comparison of online and face-to-face exams according to various variables Studies on examining the opinions and attitudes of students, pre-service teachers and teachers towards online exams			
Author/s	**Year**	**Purpose**	**Study Group/ Data Source**
Afacan Adanır, G.A., İsmailova, R., Omuraliev, A. & Muhametjanova, G.	2020	To investigate students' perceptions about online exams of 2 different countries and to compare the results	370 undergraduate students in their first-year courses online Quantitative data - via a survey Qualitative data – open-ended questions
Findings: The results provided different perceptions in terms of gender, major, and experience prior to the online course. Also, when Turkish and Kyrgyz students' ideas were compared, Turkish students had more positive thoughts and stated online exams to be less stressful, more reliable and fairer.			
Odacı, M.M.	2019	To reveal the attitudes of students who are subjected to assessment and evaluation in online exam environment through computer-based testing (CBT) platform and environment, and to examine the cognitive load of students in the exam	Consists of 32 students Case study design; English achievement test, a rating scale, a computer scale & qualitative data from the interviews of 7 volunteers.
Findings: The results exhibited that the online exam environment affected participants' attitudes and cognitive levels positively.			
Urgun, A.	2019	To compare the psychometric properties of the tests and student performances by applying three different achievement tests in a mobile (online) environment and by traditional method.	Consists of 100 high school students (Female=50, Male=50) 20-item multiple-choice achievement tests (in mobile and paper-based form)
Findings: The results provided no significant difference in terms of gender in both environments. However, in the Biology and Foreign Language (German) achievement test, a difference in favour of girls was detected.			

Author/s	Year	Purpose	Study Group/ Data Source
Pamukçu, A.	2018	To explore the contributions of online tests on success in in-class exams and students' perceptions toward them.	Quantitative data: assessment records of the repeating students (N=255) (having taken both assessed online exams and unassessed online practice exams) qualitative data: semi-structured interviews (N: 11)
Findings: The results of the quantitative data revealed important moderate beneficial correlations between formative and summative in-class assessments. Besides, qualitative data showed positive opinions of the students towards blended learning environment, the correlation of the content of the online tests and in-class tests, the effectiveness of the online tools in terms of getting ready for in-class exams and the contribution of the immediate feedback to students' success. The possible contributions of online tests to in-class success were mentioned and the need for precautions to prevent cheating in online settings was reminded.			
Aksoy, H.	2018	To introduce an application of distance education centre examination system (UZEMSS)	Following the steps for developing a web-based software
Findings: The developed online exam system was introduced and purposed to be utilized for distance education centres.			
Çörekçioğlu, S.	2017	To determine teacher perceptions and student attitudes towards the online English exams	10 Turkish and foreign teachers; 89 5th grade middle school students
Findings: The results revealed that both parties, teachers and students possess positive opinions about the use of online English exams. The advantages mentioned about online exams were time-efficient, environmentally friendly, having validity, being motivating, and giving feedback reports. Internet connection and computers were the most mentioned problems.			
Gündoğmuş, İ.	2017	To examine the assessment invariance of the paper-pencil, computer-based and tablet-based exams; to investigate whether there is any difference in exam duration and to reveal student opinions about the online application environment.	Mathematics Achievement Test with 20 multiple-choice items applied to 419 students. Semi-structured interviews with 261 students
Findings: The results stated that only configural invariance across the groups was ensured and making a comparison between the scores of the students in the groups was not valid.			

Author/s	Year	Purpose	Study Group/ Data Source
Özturan, T.	2016	To find out the impact of computer-assisted assessment on exam success and attitudes of prospective English teachers	97 pre-service English teachers who were divided into two groups: experimental and control group. The experimental group took the midterm exam in a computer environment, while the control group took the traditional method.
Findings: The results indicated that although the control group did not show any interest, a positive attitude towards computer-based exams from the experimental group was observed and the latter stated that they would be more successful through computer-based exams. Students being better equipped with computer skills and students having higher academic achievement scores used the system more easily.			
Still, M. L. & Still, J. D.	2015	To compare student learning outcomes associated with traditional in-class exams and frequent online exams	Students (N = 139) - taught by the same instructor
Findings: The results yielded no significant difference in the researched areas. However, frequent online assessments were suggested in that they may lead to positive subjective results for students.			
Candrlic, S., Ktic, M.A. & Dlab M.H.	2014	A comparative research of paper-based testing versus online testing using MudRi.	Includes three courses and 1231 tests.
Findings: The results of the research showed that the exam results of the online tests were significantly different when compared with the paper-based exams.			
Fask, A., Englander, F. & Wang, Z.	2014	To assess the difference in performance between students taking a traditional, proctored exam and those taking an online, unproctored exam.	regression models examines 44 undergraduate statistics students
Findings: The results of the study showed that the online exam testing environment created a disadvantage for the students.			
Jeong, H.	2014	To compare the scores of Korean students on computer-based and paper-based versions of the same test	Both tests included the same 80 test questions - multiple Choice

Author/s	Year	Purpose	Study Group/ Data Source
Findings: The results indicated no great difference in the scores of the two different methods which clarified that being familiar with information technology does not guarantee to adapt to computer-based tests easily.			
Jamil, M., Tariq, R.H. & Shami, P.A.	2012	To reveal teachers' perceptions about computer-based (CB) versus paper-based (PB) exams	Questionnaires - 314 teachers (To assess the validity, the instrument was piloted among 5 randomly selected teachers)
Findings: The results indicated that teachers in the study possess positive attitudes towards computer-based exams. However, in certain situations, they still prefer paper-based exams. In addition, teachers who are more equipped with computer skills are more positive for computer-based exams.			
Navruz, M.	2011	To examine the effects of computer-based tests on academic achievement	80 8th grade middle-school students were divided into two groups: taking an online test and a paper-based test
Findings: It was found that there was no statistically significant difference between classical paper-pencil tests and computer-based online tests in terms of student scores.			
Anakwe, B.	2008	To determine the impact of assessment methods on student performance and whether the use of computer-based tests instead of paper-based tests affects students' traditional test scores	75 students enrolled in any of three specific undergraduate accounting courses
Findings: The results indicated no significant difference in students' performance when computer-based and paper-based tests were compared.			
Bayazıt, A.	2007	To find out the testing time and performance differences between online and paper-based tests	46 3rd grade university students half of whom took an online test, the rest took a paper-based test
Findings: The results showed that there is not valid variation between the scores from the paper-pencil exam and the online exam. Yet, students who took the online exam needed more time. While students stated that the online exam was easy to use and made them feel comfortable, they complained about the distractions like noise, being tiring and the difficulty they experienced in focusing on the online exam.			
Wallace, P. & Clariana, R.B.	2005	To examine student performance in computer-administered and paper-based tests	207 undergraduate-level freshman business majors
Findings: The results of the study demonstrated that the scores of the students on computer-administered tests were slightly higher. While male participants outscored females in the first test, the final exam results were visa-versa.			

Author/s	Year	Purpose	Study Group/ Data Source
Campton, P.	2004	To reflect on the effectiveness of the change from a paper-based assessment to an online system, and also on the advantages and disadvantages of the change for students, lecturers and tutors	The performance of students in undertaking the unit in Semester 2, 2002 (paper-based system) and Semester 2, 2003 (online system) was measured by using a common assessment test
Findings: The results provided a critical analysis of online assessment methods and showed no significant difference between the groups in regards to performance. Several administrative benefits resulting from the online assessment method were mentioned.			
Studies on the effect of online exams on students' exam performance, anxiety and stress			
Stowell, J.R. & Bennett, D.	2010	To examine the effect of online practice exams on student performance and test anxiety	69 participants from a psychology course
Findings: The results showed that students, with high anxiety levels in classrooms settings, experienced less anxiety in online exams; on the other hand, those with low anxiety levels experienced the reverse. Additionally, there was less relationship between test anxiety and exam performance.			
Cassady, J.C. & Gridley, B.	2005	To scrutinise the effects of online formative and summative assessment on test anxiety and performance	84 undergraduate students
Findings: The results revealed some benefits for using online practice tests prior to graded course exams. Also, students stated having less anxiety while taking the tests online, and when online practice tests are integrated, they may assist students with the preparation for course exams.			
Özel, S.	2006	To reveal the predictive power of internet-based assessment applications on achievement scores	11 courses taught under the graduate program of engineering
Findings: The results manifested that according to the groups that were taking 9 courses online, there were significant stepwise regressions. Yet, this could not be viewed for the remaining 2 courses. In addition, final exam scores were affected by prediction variables at a lower level.			
Studies on the development of online exam proctoring systems/the impact of exam proctoring			
Vazquez, J. J., Chiang, E. P., & Sarmiento-Barbieri, I.	2021	To reveal the effects of proctoring methods (face-to-face and web-based) on exam scores	Two classes: face-to-face and online

Findings: The results showed that the scores of unproctored exams were 11% higher on average. Besides, the difference of scores (of proctored and unproctored exams) become greater in face-to-face exam practices. Therefore, when compared with web-based exams, proctors in face-to-face exams have a larger influence on the scores.

Author/s	Year	Purpose	Study Group/ Data Source
Jia, J. & He, Y.	2021	To design and implement an intelligent online proctoring system by using the advantage of artificial intelligence technology	The intelligent online proctoring system (IOPS) is a system that stores the identification data of all examinees and their important behaviour change status, including facial expression, eye and mouth movement and speech.

Findings: The results revealed that all participants showed full concentration during the exam and were in the camera angle throughout the exam time. However, some of the background noises could not be recorded due to technical causes.

Hylton, K., Levy, Y. & Lauire, P.D.	2016	To investigate the deterrent effect of Webcam-based proctoring on misconduct during online exams	An experimental and control group; One group monitored by a Web-based proctor; the other not monitored

Findings: The results indicated no statistically significant difference between the scores of the two groups, although the non-proctored group had slightly higher scores. However, a significant difference was found in the time taken to complete the online exams. The results of a post-experiment survey indicated that those who were not proctored were perceived to have experienced greater levels of opportunity to engage in misconduct than those who were monitored by a web-based proctor.

Harmon, O.R. & Lambrinos, J.	2008	To estimate a model that predicts exam scores from independent variables of student characteristics	62 students from two online courses - an identical exam was administered in a proctored and an unproctored setting

Findings: The results clarified that when the exams were not proctored, cheating was taking place.

Studies on the development and testing of online exam/scoring software			
Author/s	Year	Purpose	Study Group/ Data Source
Yılmaz-İnce, E.	2016	To introduce web-based automated Turkish Essay Scoring System (TSPS) software to store exam	a case study: 10 open-ended questions of a related course – 41 students

		questions and to make online exams for scoring short-answer question essays automatically.	**to ensure validity and reliability, 2 instructors created different answer keys.
Findings: The results yielded from the case study indicated that with a 92% success rate, the TSPS software can be benefitted for automated essay scoring in Turkish.				
Yağcı, M.	2012	To design a new online examination model and to make a comparison with paper-based test	70 vocational school students
Findings: The results indicated that the online exam system developed in this study enables the exam application, assessment and statistics of results which take instructors a lot of time to do manually.			
Jung, I.Y. & Yeom, H.Y.	2009	To propose a secure online exam management environment mediated by group cryptography using remote monitoring and control of ports and input	Following the step for developing a web-based software
Findings: An enhanced secure online exam management environment was proposed in this study. The system provided a solution to the most common problems of online exams, security and cheating.			
İçten, T.	2006	To develop a web-based online exam software for students taking courses online	Steps followed to develop the software
Findings: The web-based software for online exams developed in this research is capable of presenting the statistics of the exam results separately as student-based, subject-based and quality of the questions is provided by the software.			
Emir, Ş.	2006	To give information about e-Learning and to develop a web-based online exam management software	Following the step for developing a web-based software
Findings: The software will be beneficial because it is time-efficient in terms of exam application and assessment. Additionally, different methods and assessment-evaluation processes can be utilized and it is possible to follow student progress through this software.			
Çelik, Z.	2006	To develop a web-based automation and assessment evaluation management system	Steps followed to develop the software
Findings: The web-based automation and assessment-evaluation management system developed within this research provided convenience and increased students' motivation and self-confidence.			

CHAPTER III

METHODOLOGY

3.1. Introduction

In this chapter, the research methodology is presented and based on the purpose of the study, the research design, study group, data collection tools, data collection and data analysis processes are explained in detail.

3.2. Research Design

This study aimed at examining online testing practices in foreign language testing through the lens of instructors and students. The study was planned and conducted in accordance with the principles of the case study design, one of the qualitative research methods. A case study is a methodological approach that involves an in-depth study of a restricted system using multiple data collection to gather systematic information about how it works (Chmiliar, 2010). In case studies, the researcher examines one situation or a few situations limited in time with data collection tools containing multiple sources (observations, interviews, audio-visuals, documents and reports) defines the situations and themes depending on the situation (Creswell, 2007) and seek answers to "how" and why" questions (Yin, 2009). Unlike experimental studies, the researcher conducting the case study does not attempt to compare but to explore and tries to define categories of events and behaviours instead of testing the hypothesis or proving relationships (Hancock & Algozzine, 2006). The use of case studies explains, describes, and discovers events that are supposed to have causal links and involve too many interventions that cannot be explained by experimental or survey methods in real life and distinguishes it from other studies (Yin, 1994).

In the literature, the case studies were classified under the following headings: single case-holistic design, single-case embedded design, multiple case-holistic design, multiple case-embedded design (Yin, 1994); exploratory, explanatory, descriptive (Yin, 1994); disciplinary orientation, overall intent, multiple case studies (Merriam, 1998); theory-seeking and theory testing, story-telling and picture-drawing, evaluative (Bassey, 1999); intrinsic, instrumental,

collective (Stake, 2005). Since the objectives of the study are twofold: in terms of instructors and students, the multiple case-holistic design proposed by Yin (1984) was adopted in this research. In this design, each case is considered holistically on its own, then compared. In this study, the opinions of instructors and students on online testing in foreign language teaching are discussed, the findings are compared to reach conclusions.

In the research, coding method was used to ensure the confidentiality of the participants' identities. Accordingly, students' identities were coded as S+Gender (F/M)+ Number (1,2,3...). The coding: SF13 in the research findings indicates that this a female student and she is in the 13th line. Similarly, the identities of the instructors were coded as I (Instructor) + F/M (Female/Male) + Number (1,2,3...). The code: IF15 in the research findings expresses this is a female instructor and she is in the 15th line.

3.3. Study Group

The study group consisted of 26 instructors working at a private university located in a large province in Southern Turkey and 134 students studying in the preparatory school of the same university during the spring semester of the 2020-2021 academic year. The participants of the research were determined through the convenience sampling technique to bring speed and practicality to the research. On the other hand, the fact that the participants are close and accessible to the researcher has also been effective in sample selection (Yıldırım & Şimşek, 2018) regarding the COVID-19 pandemic conditions. In addition, criterion sampling, one of the purposeful sampling methods, was used to determine the sample. In this context, the basic criteria for the selection of the participants to be included in the study are that instructors have proctored an online and students have similarly taken the online exam for at least one academic year. At the university where the study was conducted, data collection tools were sent to English instructors and preparatory students who met the relevant criteria and were asked to participate in the study on a voluntary basis. The computer literacy data are self-related, no assessment to test the data is done within the research.

Table 3. 1. *Demographic statistics for students*

Variable	Sub-level	*f*	%
Gender	Female	79	58,96
	Male	55	41,04
Age	18	38	28,36
	19	45	33,58
	20	25	18,66
	21	13	9,70
	22 and above	13	9,70
English Level	Elementary	7	5,22
	Pre-intermediate	40	29,85
	Intermediate	41	30,60
	Upper-intermediate	23	17,16
	Advanced	23	17,16
Computer Skills	Basic	18	13,43
	Medium	70	52,24
	Advanced	39	29,10
	Expert	73	54,48
Internet Usage Background	1-3 years	12	8,96
	3-5 years	17	12,69
	More than 5 years	105	78,36

Table 3. 2. *Demographic statistics for instructors*

Variable	Sub-level	*f*	%
Gender	Female	23	95,00
	Male	3	5,00
Age	25-30	11	35,00
	32-37	9	35,00
	40 and above	6	30,00
Educational Background	Bachelor's degree	15	45,00
	Master of Arts	11	55,00
Teaching Experience in the Institution	6 months-3 years	13	60,00
	4-6 years	10	30,00
	8 years	3	10,00
Teaching Experience (Overall)	1-4 years	4	10,00
	5-10 years	13	45,00
	More than 10 years	9	45,00
Computer Skills	Basic	2	10,00
	Medium	17	70,00
	Advanced	7	20,00
Active Internet Usage Background	Less than one year	1	5,00
	5-10 years	7	30,00
	More than 10 years	18	65,00
Distance/online Education Experience	Yes	5	
	No	21	

3.3.1. Setting and Online Exam Details

The current study is carried out in a private university located in the south of Turkey. In this university the exams are held in two sessions; in the first session reading, listening and writing skills are assessed respectively, and speaking is assessed in the second session. Throughout the exams, all students are required to keep their cameras and microphones on, except for the listening part in which merely cameras are on. All the questions in the exams are prepared in the form of open-ended since the university encourages students to advance their communicative skills. The steps below are followed during the exams:

The reading part lasts for 45 minutes and it consists of two different reading tasks with five questions in each. All students in a class take this part together, yet they complete their exam in separate rooms (in each room the maximum number of students is nine). When the time finishes, the students stop writing and take a photo of their papers in order to send it to the proctor via email.

The listening part of the first session lasts 30 minutes and the students in the class are divided into two. In this section, the students have two different listening tasks. The first task is while listening. Students listen to the recording and answer five questions related to the listening. Thereafter, students are to take the lecture part of the exam. Firstly, students listen to the task two times and take notes. Then they get the questions of the lecture and are given ten minutes to answer them. At the end, they send their papers to the proctor.

Writing part is 60 minutes at length. Before the exam time starts, students are given time to read and take notes of the questions as they are not allowed to look at the screen during the exam time. This is because of the cheating concerns. Same as the reading part, they take their exam in different rooms with a maximum of nine people in each. They send their papers to the proctor in the end.

The second session, speaking, is individual and assessed by two raters. The speaking exam consists of two questions, and the student needs to answers them separately, each answer needs to last for two minutes. If the student does not want to answer the question, s/he can change only one of the questions. For each question, the student is given one minute of thinking and note-taking time. When needed the student is provided with follow-up questions.

3.4. Data Collection Tools

Two data collection tools were used to collect the data proceeded in the study. The tools consist of two online structured opinion forms prepared by the researcher. After the relevant

literature review on the concepts like testing, online testing, and testing-evaluation in foreign language education, a question pool was created for both forms. Later, the questions that will best serve the purpose of the research were determined by the researcher and the supervisor. Prior to administering the forms, the questions in the final forms were sent to two field experts who had the degree of Doctor of Philosophy (PhD.) in English Language Teaching (ELT) and are currently carrying out studies on testing and evaluation in foreign language education and two Turkish language experts to ensure the content validity of the forms. After the questions in the forms were rearranged in line with the experts' feedbacks, both forms were piloted to test the functionality of the forms and minor changes were made in the questions to increase the understandability. Whereas the form prepared for students was administered in Turkish so that they could express their opinions and thoughts easily, the other was conducted in English. To avoid endangering the health of both researchers and participants from the Covid-19 pandemic that has claimed millions of people worldwide since it was first reported in November 2019, both forms were distributed to the participants online through the platforms specifically designed for this purpose.

3.4.1. Online Instructor Opinion Form

The first data collection tool is the online instructor opinion form encompassing seven demographic questions and six open-ended questions that lead the instructors to express their opinions on online testing in foreign language teaching from different perspectives (see Appendix A). The questions mainly focused on the online testing practices in the participants' institution and their practicality, deficiency/deficiencies they have experienced in online testing, the participants' potential suggestions for the deficiencies they have detected in online testing, the presence of problematic skill/s in online testing, common problems in online testing, advantages/disadvantages of online testing and alternative model suggestions to current online practices.

3.4.2. Online Student Opinion Form

The second data collection tool is the online student opinion form consisting of three demographic questions, six multiple-choice and nine open-ended questions about online testing in foreign language teaching (see Appendix B). In this form, the participants were firstly asked to express their opinions by scoring and selecting choices. In the multiple-choice questions part, the participants' self-evaluation on success and stress status on face-to-face and online exams, computer skills and internet usage background were asked. In the open-ended questions part,

they were asked to express opinions regarding the advantages/disadvantages of online testing, general problems they had in online exams, sections that need changing/improving, time management in online exams, the effect of online exams on their exam performance, the hardest skill/s in online exams, source of problems encountered in online exams, text type choices, and skill/s preference in case of maintaining online testing.

3.5. Data Collection Process

The data collection process started by getting the relevant permissions from the authorities. First, the ethics committee permission form and research application permission petition were filled. After the approval of the relevant documents, the legal permit procedures were completed. Considering the COVID-19 pandemic, the data collection tools planned to be implemented online were distributed to the participants after they were loaded into the system where the application would be made. The forms were reached to the participants through their e-mail and social media addresses, and they were asked to fill in the relevant forms.

3.6. Data Analysis

Analysis of the data obtained in a case study depends on making a detailed description of the situation and the environment. Therefore, data analysis was first started by describing the environment and explaining which processes were followed. The data analysis process was first started with quantitative data. A commonly used quantitative data statistical package program was used in the analysis of quantitative data collected from students. Data were tested for normality to decide on the quantitative data analysis to be applied. Normality tests were performed based on the three criteria. The first of these was the Kolmogorov-Smirnov (K-S) and Shapiro Wilk test as a statistical hypothesis approach. These tests perform analyses on the null hypothesis that the data group has a normal distribution. If the significance value obtained as a result of the analysis is greater than .05, the data is normally distributed, or else, it deviates from the normal (McKillup, 2012). The 'Z test', which is calculated as a result of dividing the skewness/kurtosis value in its own standard error, was another criteria used for this purpose. If the "Z" value is less than 1.96, it provides evidence that the data are normally distributed (Field, 2009; Howitt & Cramer, 2011). Another criterion was checking the skewness and kurtosis values for the normality test. If the relevant values are within the tolerance limit of +/-1, the data are considered to be normally distributed (Büyüköztürk, 2016). As a result of the normality test, non-parametric data analyses were used as deviations from normal were observed in the

examined groups. In order to make comparisons between the groups, Mann Whitney U Test was used for the variables consisting of two sub-levels and Kruskal Wallis H Test was used for the variables consisting of three or more sub-levels.

In the second phase, the qualitative data collected in the research were analysed. The content analysis technique was used to analyse them. The content analysis is to classify, to convert into numbers and/or to make inferences objectively and systematically about a message in verbal, written and other materials (Tavşancıl & Aslan, 2001). By examining the data, the researcher tried to divide it into meaningful sections and to understand what each section means conceptually. These parts, which form a meaningful whole in themselves, were named by the researcher.
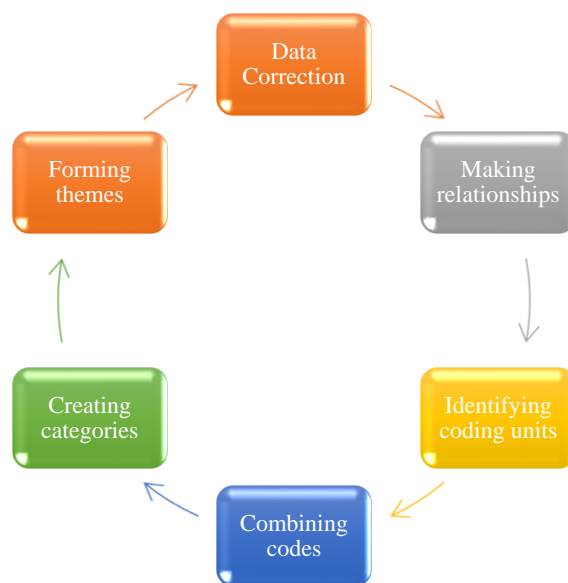
Figure 3.1. *Reporting stages of content analysis*

The qualitative data analysis process was conducted by two experts following the sequence in figure 3.1. It was started with the data correction process in which some participants whose statements did not serve the objective of the research were eliminated, the data obtained from the participants were downloaded from the platform and tailored to the data analysis template. Later, it was moved to the further stage where it was tried to make relationships among the participants' statements. Based on the relationships, codes were identified in the following stage. Since the codes were the smallest units having relationships, they were combined to reach the larger units "categories".   In line with research questions, the obtained categories were grouped under the themes.

Since two experts took part in the analysis process, Kappa analysis was conducted to measure inter-rater reliability or agreement between the experts. The Kappa analysis can be calculated in two ways, namely Cohen and Fleiss Kappa. While measuring the agreement

between two reviewers, Cohen Kappa; when the number of reviewers is more than two, Fleiss Kappa analysis is used to measure inter-rater reliability (Kılıç, 2015). Since two experts interpreted the data collected within the scope of the research, Cohen Kappa analysis was performed to check the inter-rater reliability. As a result of the analysis, the Kappa value between the experts was calculated as .83. According to reference intervals (see Table 3), provided by Landis and Koch (1977) "almost perfect" level of agreement was observed among the experts. After the content analysis was completed, the findings were presented in tables together with the participant statements.

Table 3. 3. *Reference intervals for kappa analysis*

Interval	Level of agreement
.01--.20	Slight
.21--.40	Fair
.41--.60	Moderate
.61-- .80	Substantial
.81--1.00	Almost perfect

CHAPTER IV

FINDINGS

In this chapter, the findings obtained as a result of the examination of the data collected within the scope of the research are presented with the support of tables, figures and participant statements.

4.1. Findings for the First Research Question

In this part, the findings related to the first research question were presented. In the related research question, it was examined whether the students' success and stress levels in online and face-to-face exams differed according to their demographic characteristics (gender, age, English level and computer skills). The findings obtained as a result of the analysis are presented in Table 4.1, Table 4.2, Table 4.3, Table 4.4 and Table 4.5.

Table 4. 1. *Students' success and stress situations in face-to-face and online exams*

Dimension	N	Min.	Maks.	X̄/SD	Mode
Success in face-to-face exams	134	2	5	3,84±0,76	4
Success in online exams	134	1	5	3,66±0,91	4
Stress in face-to-face exams	134	1	5	3,46±1,35	4
Stress in online exams	134	1	5	3,22±1,31	3

In Table 4.1, descriptive statistics of students' beliefs on success and stress levels in face-to-face and online exams were presented. Accordingly, students underlined a high level of success both in face-to-face and online exams. Regarding the mode of success beliefs in face-to-face and online exams, they may be regarded as equal; however, it was seen they believed to show higher success in face-to-face exams compared with online exams considering the mean values. When the students' views on stress levels in face-to-face and online exams were examined, it was observed that they had greater stress in face-to-face exams compared to online exams. Both mean values and modes in the stress dimension were the indicators of this finding.

Table 4. 2. *Students' success and stress situations in face-to-face and online exams (by gender)*

	Gender	N	Mean ranks	Sum of Ranks	U	p
Success in face-to-face exams	Female	79	66,58	5259,50	2099,500	0,722
	Male	55	68,83	3785,50		
Success in online exams	Female	79	65,92	5208,00	2048,000	0,550
	Male	55	69,76	3837,00		
Stress in face-to-face exams	Female	79	75,22	5942,00	1563,000	0,005
	Male	55	56,42	3103,00		
Stress in online exams	Female	79	73,75	5826,00	1679,000	0,022
	Male	55	58,53	3219,00		

Table 4.2 presents the comparison results of students' success and stress levels in face-to-face and online exams by gender. According to the analysis results, it was found that students 'views on success level in face-to-face (U=2099.500; p>.05) and online exams (U=2048.000; p>.05) did not significantly differ based on their genders. On the other hand, students' views on stress levels in face-to-face (U=1563.000; p<.05) and online exams (U=1679.000; p<.05) significantly changed according to their genders. When the mean ranks were examined, it was seen that female students (75.22; 73.75) felt higher stress both in face-to-face and online exams than males (56.42; 58.53).

Table 4. 3. *Students' success and stress situations in face-to-face and online exams (by age)*

Dimension	Age	N	Mean Rank	df	χ^2	p	Difference*
Success in face-to-face exams	18 [1]	38	71,86	3	10,537	0,015	1-2,1-4
	19 [2]	45	54,53				
	20 [3]	25	65,46				
	21 [4]	13	43,08				
Success in online exams	18 [1]	38	64,11	3	2,912	0,405	-
	19 [2]	45	54,43				
	20 [3]	25	64,98				
	21 [4]	13	67,00				
Stress in face-to-face exams	18 [1]	38	56,61	3	3,597	0,308	-
	19 [2]	45	60,12				
	20 [3]	25	60,88				
	21 [4]	13	77,12				
Stress in online exams	18 [1]	38	64,16	3	1,583	0,663	-
	19 [2]	45	61,37				
	20 [3]	25	53,74				
	21 [4]	13	64,46				

In Table 4.3, comparison results of students 'views on success and stress levels in face-to-face and online exams by age were presented. Accordingly, students' views on success levels in face-to-face exams significantly differed in terms of their ages (χ^2=10.537; p< .05). As a result of the Mann Whitney Test conducted to determine between which groups there was significant differentiation, it was revealed there was a significant difference between group 1-group 2 and group 1-group 4. Mean ranks showed that 18-aged students believed to have higher success in face-to-face exams compared to 19 and 21-aged students. In other words, it can be implied that the older students were, the lower their beliefs on success in face-to-face exams would be. On the other hands, students' views on success in online exams (χ^2=2.912; p> .05), and stress in face-to-face (χ^2=3.597; p> .05) and online exams (χ^2=1.583; p> .05) did not significantly change by age.

Table 4. 4. *Students' success and stress situations in face-to-face and online exams (by level)*

Dimension	Level	N	Mean Rank	df	χ^2	p	Difference*
Success in face-to-face exams	Elementary [1]	7	96,07	4	8,296	0,081	
	Pre-intermediate [2]	40	70,10				
	Intermediate [3]	41	57,74				
	Upper-intermediate [4]	23	65,85				
	Advanced [5]	23	73,33				
Success in online exams	Elementary [1]	7	81,14	4	19,674	0,001	2-3,2-4,2-5
	Pre-intermediate [2]	40	46,40				
	Intermediate [3]	41	77,02				
	Upper-intermediate [4]	23	79,30				
	Advanced [5]	23	71,26				
Stress in face-to-face exams	Elementary [1]	7	95,50	4	7,418	0,115	-
	Pre-intermediate [2]	40	63,41				
	Intermediate [3]	41	73,13				
	Upper-intermediate [4]	23	67,63				
	Advanced [5]	23	55,91				
Stress in online exams	Elementary [1]	7	46,86	4	4,893	0,298	-
	Pre-intermediate [2]	40	75,19				
	Intermediate [3]	41	65,07				
	Upper-intermediate [4]	23	60,67				
	Advanced [5]	23	71,57				

In Table 4.4, comparison results of students 'views on success and stress levels in face-to-face and online exams' by English level were presented. As seen in the table, while students' views on success level in online exams significantly changed according to their levels in English, their views on success in face-to-face exams and stress level in face-to-face and online exams did not meaningfully differ by English level. In order to find the binary group where there was a significant difference, Mann Whitney Test was conducted in the dimension of success in online exams. The analysis results showed that there was significant differentiation between the groups 2-3, 2-4 and 2-5. When the mean ranks were examined, it was revealed that students in the pre-intermediate level had lower success belief in online exams than the students in intermediate, upper-intermediate and advanced levels. In other words, as from pre-intermediate level, as the students' English level increased, their views on success in online exams decreases.

Table 4. 5. *Students' success and stress situations in face-to-face and online exams (by computer skills)*

Dimension	Computer skill	N	Mean Rank	df	χ^2	p	Difference*
Success in face-to-face exams	Novice [1]	18	77,69	3	4,784	0,188	-
	Medium [2]	70	63,44				
	Advanced [3]	39	66,29				
	Expert [4]	7	88,57				
Success in online exams	Novice [1]	18	44,97	3	33,340	0,000	1-3
	Medium [2]	70	57,10				1-4
	Advanced [3]	39	91,77				2-3
	Expert [4]	7	94,21				2-4
Stress in face-to-face exams	Novice [1]	18	75,56	3	3,734	0,292	-
	Medium [2]	70	62,57				
	Advanced [3]	39	69,37				
	Expert [4]	7	85,64				
Stress in online exams	Novice [1]	18	95,53	3	30,894	0,000	1-2
	Medium [2]	70	75,59				1-3
	Advanced [3]	39	42,46				1-4
	Expert [4]	7	54,00				2-3

In Table 4.5, comparison results of students 'views on success and stress levels in face-to-face and online exams by computer skills were presented. As seen in the table, while students' views on success in online exams did not significantly differ by computer skills (χ^2=4.784; p> .05), their views on success in online exams significantly changed according to their computer skills (χ^2=33.340; p< .05). As a result of the Mann Whitney Test, which was conducted to find out in which groups there was differentiation, it was determined that there was a significant differentiation between the groups 1-3, 1-4,2-3 and 2-4. The mean ranks showed that students with both novice and medium level computer skills had lower success belief in online exams than the students with both advanced and expert computer skills. According to analysis results, students' views on stress in face-to-face exams did not significantly change by computer skills. On the contrary, their views on stress in online exams significantly changed according to their computer skills. As a result of the Mann Whitney Test, which was conducted to find out in which groups there was differentiation, it was determined that there was a significant differentiation between the groups 1-2, 1-3, 1-4 and 2-3. The mean ranks revealed that the students with novice computer skills had greater stress in online exams than those with medium, advanced and expert computer skills. Besides, the students with medium computer skills had higher stress in online exams than the students with advanced computer skills.

4.2. Findings for the Second Research Question

In this part, the findings related to the second research question "What are the students' views on online testing practices in foreign language education?" were presented. In the related research question, students' views on online exams were examined from various aspects. Accordingly, the students' views on advantages and disadvantages of online exams (Table 4.6), general problems they encountered in online exams (Table 4.7), the aspects of online exams that need to be changed/improved (Table 4.8), whether additional time should be given in online exams (Table 4.9), whether they can show their real performance in online exams (Table 4.10), the skill/skills they have the most difficulty in online exams (Table 4.11), whether the problems they encounter in online exams are due to the administration way of exams or general problems (Table 4.12), the preferences for exam types (online or face-to-face) (Table 4.13) and the skill/skills they wish to continue with online exams (Table 4.14) were examined and the findings were presented together with some selected statements of the students.

Table 4.6. *Students' views on advantages and disadvantages of online exams*

	Theme/Code	Number of Participants	Total Number of Participants
Advantages	Less stressful exam atmosphere	40	
	Flexibility of place	16	
	Better concentration	3	
Disadvantages	Connection/technical issues	46	
	Stressful exam procedure	15	107
	Cheating conducive environment	9	
	Adaptation/concentration problem	5	
	Communication problem	1	
	Difficulty in finding a convenient place	1	

Table 4.6 presents the findings of students' views on the advantages and disadvantages of online exams. As a result of examining the collected data with content analysis technique; three codes respectively "less stressful exam atmosphere, the flexibility of place and better concentration" for advantage and six codes namely "connection/technical issues, stressful exam procedure, cheating conductive environment, adaptation/concentration problem, communication problem and difficulty in finding a convenient place" for disadvantages emerged. The code "less stressful exam atmosphere" is the most expressed code (f=40) by the students in the advantages theme. Through this code, the students emphasized the advantage of online exams to face-to-face exams in terms of stress dimension and stated to have lesser stress in online exams. The second highest expressed code was the flexibility of place in this theme (f=16). In this code, the students pointed out the nature of online exams and underlined the place convenience of online exams. The last code emerged in the advantages theme was better concentration. The students marked that their concentration advanced in online exams compared to face-to-face counterparts. In terms of disadvantages, the connection/technical issues code was the highest expressed among the students. Through this code, the students stated that although online exams had some appealing advantages for them, connection/technical issues were the greatest drawbacks of these exams. The second most expressed code was stressful exam procedure in this theme. Although some students stated online exams created a stress-free environment, others were not on the same line and specified

that online exams caused stressful exam procedures. The third code was cheating conductive environment. Some students complained about the cheating conductive aspect of online exams and stated that some of their classmates had higher marks in the exams due to this handicap of the exams. The fourth code was the adaptation/concentration problem. Through this code, the students underlined that they had adaptation/concentration problems in online exams. The problems generally occurred since their classmates' microphones were on during the exams and there were constantly sounds. Some students added they were not used to taking exams online, so they had difficulty at that point. Communication problems and difficulty in finding a convenient place were the other codes expressed in this theme. In these codes, students expressed that due to the quality of internet connection, they had some communication problems, especially while interacting with instructors. Besides, one of the students drew attention to the convenience of the place. Since the exams were online, they occasionally had difficulty in finding an appropriate place to take the exams.

Some of the students' statements on this theme were as follows:

Advantages

I think that the stress level in exams is reduced in the environment we are used to. (SM13, less stressful exam atmosphere)

Internet and connection problems are a disadvantage, but face-to-face exams are more stressful. Convenience is an advantage since the questions that can be asked in online exams are limited. (SF41, less stressful exam atmosphere)

You do not have to leave the house, you can take the exam from where you are sitting. (SF6, Flexibility of place)

I can go to class as soon as I wake up. (SF106, Flexibility of place)

I can concentrate better since I am alone. (SF36, Better concentration)

It allows me to focus only on the exam without experiencing stress dependent on anything other than the exam (road, clothes, weather, etc.) (SM48, Better concentration)

Disadvantages

I think that the computer or internet problem in online exams affects the student badly. (SF2, Connection/technical issues)

It seems more comfortable because we are at home, but it is more stressful and we continue with the fear that my exam will be closed in the middle of the exam because our internet connection is not perfect and we can't get a good computer because of the computer prices. (SM85, Connection/technical issues)

I think it has more disadvantages. A person becomes more stressed during the exam and becomes more open to making mistakes. (SF33, Stressful exam procedure)

We are more likely to be stressed in an online exam. (SF66, Stressful exam procedure)

The opportunity of cheating is too much, there is a possibility that there is any problem with the internet or computer. (SF5, Cheating conducive environment)

For me, online or face-to-face exams are almost the same, but I think the online cheating rate is higher. (SM8, Cheating conducive environment)

Since all our academic studies and evaluations are on a single device, many of our works are progressing more regularly, but due to the current conditions, the online system and the fact that we are forced to take the exam in our private area prevents us from adapting to the exam. (SF1, Adaptation/concentration problem)

I think there is no advantage, on the contrary, it has a disadvantage because it is very difficult to focus on the exam on a computer screen, and at the same time, everyone, including myself, can perform better in face-to-face exams than online. (SF14, Adaptation/concentration problem)

We may have problems with not understanding the exam fully. (SM23, Communication problem)

The environment where we will take the exam may not be suitable. (SM73, Difficulty in finding a convenient place)

Table 4. 7. *General problems students encountered in online exams*

Problems	Number of participants	Total number of participants
Connection/technical problems	76	
Sound/noise problems	11	
Stress	5	93
Camera problems	4	
Cheating attempts	1	
Insufficient exam time	1	

In Table 4.7, students' views on general problems encountered in online exams were presented. As seen in the table, the problems were grouped under six codes. The first and most expressed code was "connection/technical problems". A significant number of students complained about this problem and considered it as the biggest handicap. The second most expressed code was sound/noise problems. According to the students, this problem occurred due to one of the requirements of the exams. The students stated that since they were asked to

keep microphones on, it caused some noises either from each other or devices; so they had to cope with adaptation and concentration problems as well. The third problem they encountered in online exams was stress. While some students underlined they had stress in face-to-face exams, others pointed out the stress emerged in online exams. When the students' statements were examined, it was observed that they had stress when they think as 'what if I come across an internet disconnection problem during the exam and my paper is not accepted'. Another problem expressed in this theme was camera problems. In this code, the students stated that they had trouble adjusting the camera angle, they had concentration problems because the cameras were always on, and therefore they wished to turn off their cameras. Cheating attempts and insufficient exam time were the other codes expressed in this theme. One of the students noted their classmates attempted to cheat and therefore received undeserved high grades. Another student stated that the insufficient time given in online exams was one of the problems they faced.

Some of the students' statements on this theme were as follows:

Power outages and internet connection problems (SF129, Connection/technical problems)

Freezing due to technical problems, the sound cut off (SM131, Connection/technical problems)

Microphones have to be on, external sounds can be disturbing, also adjusting the camera angle, checking the weakening internet connection can cause a waste of time during the exam. (SF7, Sound/noise problems)

In the listening exam, the sound is muffled due to the internet and I cannot understand the audio. (SF120, Sound/noise problems)

Internet problem and stress (SM73, Stress)

Being on the record makes me nervous, and the teachers get angry when I look at the screen. (SF96, Stress)

I'm just having trouble adjusting the camera angle. (SF96, Camera problems)

Camera position adjustment and connection problem. (SF128, Camera problems)

Cheating and some of them not being noticed. (SF84, Cheating attempts)

I panicked because the time for the midterm exams was too short and I could not pass the exam. (SF30, Insufficient exam time)

Table 4. 8. *Students' views on the aspects of online exams that need to be changed/improved*

Are there aspects that need to be changed/improved in online exams?	YES		NO	
	f	%	*f*	%
	57	59,38	39	40,63

		Number of participants	Total number of participants
Aspects that need to be changed/improved in online exams	Extending the exam period	12	
	Removing the requirement for microphones to be on	9	
	Removing the requirement for cameras to be on	5	
	Conducting additional listening	4	
	Changing question types	3	
	Removing the camera angle rule	2	
	Bending the rules	2	
	Fixing technical issues	2	
	Changing exam platform	2	
	Shortening the exam time	1	50
	Conducting the speaking exam with several students rather than one-on-one	1	
	Adding grammar section	1	
	Not treating as cheating in disconnection	1	
	Increasing the degree of difficulty of the questions	1	
	Increasing the number of cameras	1	
	Ensuring the instructions are understood	1	
	Preparing speaking questions of equivalent difficulty	1	
	Omitting the listening section	1	

In Table 4.8, the students' views on the aspects of online exams that need to be changed/improved were presented. As a result of examining the students' statements, 18 codes were found in this theme. Out of these codes, extending the exam period was the most expressed code. In this code, the students suggested that the exam duration should be extended in online exams, by this way, the time they lost due to connection features should be compensated and added that this was necessary because the content of the exams was more difficult than face-to-face exams. Removing the requirement for microphones to be on was the second most expressed aspect. Like in other questions, the students complained that noises occurred during

the exam due to the microphones being turned on, and therefore they had trouble focusing on the exams. Like the previous one, removing the requirement for cameras to be on code caused similar problems in online exams. The students announced it was difficult to adjust the camera angle, so they wasted their exam time. Some students added this requirement caused them to cope with concentration problems. Another aspect that needs changing/improving was conducting additional listening. It was especially expressed by the students who stated they had difficulty in understanding the audios. The students also noted that question types should be diversified in order to advance the validity of the exams. The codes "removing the camera angle rule, bending the rules, shortening the exam time, fixing technical issues and changing exam platform" were the ones each expressed by two students. The students also suggested conducting the speaking exam with several students rather than one-on-one. By this time, they could control their anxiety and have an exam process in a conversational mood. One of the students asked to add a grammar section as well as other skills. Referring to disconnection issues, another student requested no to be treated as cheating. Considering the degree of difficulty of the question in exams as insufficient, another student remarked that questions having a higher degree of difficulty should be asked in exams. While some students complained about cameras, one of them asked the proctors to increase the number of cameras in exams, so cheating attempts could be inhibited. Besides, another student highlighted the significance of instructions by asking to ensure the instructions were understood. One of the students pointed out one-on-one speaking sessions and asked the instructors to prepare speaking questions of equivalent difficulty. The last code of this theme was omitting the listening section. In this code, the student highlighted the difficulty of maintaining listening sections due to the connection issues and believed that removing it from the exam could be one of the aspects that needs changing/ improving in online exams.

Some of the students' statements on this theme were as follows:

Based on the internet speed, the duration can be extended a little more on behalf of viewing the question. (SM25, Extending the exam period)

More time can be given and additional time can be given for situations such as power outages. (SM43, Extending the exam period)

The rules that the camera and microphone are on and headphones cannot be worn in the exams should be removed because it is not ethical to make such a request from every student, even if it is due to cheating, and at the same time, everyone stays at home and wants a suitable exam space in a crowded and noisy family environment. (SF1, Removing the requirement for microphones to be on)

I wish we didn't hear the voices of other students because as I said, the fan noise can prevent us from focusing, or the microphone of some people can be sensitive, mine is sensitive, the sound of pencils and the sound of drinking water are transmitted to us very loudly. (SF20, Removing the requirement for microphones to be on)

The camera and microphone requirement should be abolished. In my opinion, the condition of the house at that time may not be available. This leads me to abandon the exam or fail the exam. (SM86, Removing the requirement for cameras to be on)

I think the camera requirement should be removed or at least it should not be a problem for us to look at the computer screen because we may have to look for the question. (SF96, Removing the requirement for cameras to be on)

An extra listening can be given in the listening section. (SM23, Conducting additional listening)

We can have more time in the listening parts or we can listen more than 2 times. (SF97, Conducting additional listening)

The exam may be multiple choice. (SF38, Changing question types)

Multiple-choice exams via the link should definitely be abolished, as public schools do. This method is only made so that students can cheat more easily and teachers do not have to deal with it. Exams with lively open-ended questions should be preferred and if the student is right in case of potential problems, these situations should be tolerated. (SM133, Changing question types)

It is said that the computer should be put away, but then it becomes difficult to read what is written, using a phone would be very good for reading at least. (SF65, Removing the camera angle rule)

The camera angle rule should be removed. (SM74, Removing the camera angle rule)

Rules should be more flexible. (SF77, Bending the rules)

You can ease some of the rules that students have to apply during the exam. Sometimes it can be very boring and students may say to themselves, "Why am I studying prep?" For example, you say that students cannot look at the screen in the writing exam, this is a ban, and this can be annoying. Involuntarily, I can look at the screen and feel bad. It would be awesome if this was removed. (SM132, Bending the rules)

Technical problems should be fixed and one-to-one classes can be recorded… (SF34, Fixing technical issues)

Online exams must be taken from an online site. (SM47, Changing exam platform)

We should do the listening exams on a site. (SM49, Changing exam platform)

The exam time can be shortened in case of cheating. (SF3, Shortening the exam time)

The Speaking exam can be in the form of a conversation with 2-3 students. This is important for reducing stress. (SF38, Conducting the speaking exam with several students rather than one-on-one)

I think the grammar section should be added as well. (SF38, Adding grammar section)

If students have internet-cut during the exam, they should be not considered as a direct cheat. (SF55, Not treating as cheating in disconnection)

The questions can be more difficult. Questions easier than our level are asked. (SF58, Increasing the degree of difficulty of the questions)

It should be followed by a few cameras against the risk of cheating. (SM60, Increasing the number of cameras)

Make sure everyone understands all the instructions in every part. (SF6, Ensuring the instructions are understood)

I think the level of some questions in the speaking section is inconsistent. Some students receive questions that are easier to respond, while others receive more difficult questions. (SF2, Preparing speaking questions of equivalent difficulty)

There shouldn't be a listening section. Sometimes we don't understand and that's not our problem. (SF102, Omitting the listening section)

Table 4. 9. *Students' views on whether additional time should be given in online exams*

		YES		NO	
Do you think that additional time should be given for online exams?		*f*	**%**	*f*	**%**
		78	72,90	29	27,10
		Number of participants		Total number of participants	
	Difference in connection/technical features	42			
	Difficulty in concentrating on the exam	6			
	Different environmental conditions	3			
	Need for the listening section	4			
	Need for the reading section	2			
Reasons/issues for additional time	Testing platform difference	2		78	
	The difference in computer skills	2			
	Need for the writing section	2			
	Frequent warnings	2			
	High stress/excitement	2			
	Exam familiarization and device control	2			

In Table 4.9, students' views on whether additional time should be given in online exams were presented. While 72, 90 % of the students thought that additional time should be given for online exams, the rest were on the opposite side. When the statements of the students who believed that additional time should be given were examined, 11 codes were identified, giving reasons for additional time. The most expressed code among these codes was the difference in connection/technical features. Through this code, the students highlighted the difference in the quality of connections and devices they used for exams, so they believed that additional time should be allocated for online exams.

Some of the students' statements on this theme were as follows:

Yes, I think because there is a high probability of problems with the computer or the internet. (SF2, Difference in connection/technical features)

Yes, because we may have a problem with technological devices or the internet and we may lose time. (SF32, Difference in connection/technical features)

Maybe because it's harder to focus. (SF6, Difficulty in concentrating on the exam)

I definitely think so. I also feel under pressure (in case I have internet or computer problems) and cannot fully focus on the questions. (SF130, Difficulty in concentrating on the exam)

Yes, because the environmental conditions are not the same for everyone. (SF40, Different environmental conditions)

Yes, because it's harder to focus at home than in the exam room. (SM43, Different environmental conditions)

I think that 2 minutes extra in the listening exam can be more productive in terms of correcting recognizable sentences. (SF72, Need for listening section)

I really think that the time is very insufficient, especially in listening. This is my 3rd step and I have the most difficulty in listening in exams. (SF83, Need for listening section)

It could be in the reading section. (SM9, Need for reading section)

It can be given in the reading exam because it takes time to read on the phone and find the answer and write it on paper because we have two long texts. (SF127, Need for reading section)

Yes, it can be because both are not the same platform, there may be delays because it takes place through the application. (SM22, Testing platform difference)

Yes, I think because no matter how fast the computer is, doing it from many tabs or many screens takes a long time. In face-to-face exams, we only focus on our papers. On the computer, we have a lot of focus, such as the screen, paper, our teacher and of course the small but insignificant sounds in our classroom. (SM85, Testing platform difference)

Exactly. Everyone does not have the same computer skills. (SF17, Difference in computer skills)

Exactly. Everyone's ability to use a pc or internet speed may not be the same. (SF31, Difference in computer skills)

Exactly. For example, in the writing section, there are some letters that we need to fill in, even if we are given a certain time, it is not enough. (SM23, Need for writing section)

Maybe extra time can be given for the writing exam because sometimes we don't have any ideas. (SF128, Need for writing section)

In some cases, teachers give too many warnings and this wastes time. (SM54, Frequent warnings)

Yes because the stress is too much. (SM73, High stress/excitement)

Absolutely yes, because we take the exam under stress and we may have to read more than a few times, and I think the time is insufficient in sections like listening. (SF84, High stress/excitement)

An extra 10 minutes would be great because it takes at least 5 minutes for students to get used to the exam environment via the online platform. At least so do I. (SM132, Exam familiarization and device control)

Table 4.10. *Students' views on whether they can show their real performance in online exams*

		YES		NO	
Could you show your real performance in online exams? Why?		*f*	**%**	*f*	**%**
		63	59,43	31	29,25
		Number of participants		Total number of participants	
Reasons for showing real performance	Showing equal performance with face-to-face exams	10			
	Stress/anxiety-free exam procedure	5			
	Showing better performance than face-to-face exams	3		23	
	The difference is not in the exams, but in the way the exam is administered.	3			
	Performance is knowledge-based, not platform-based	2			
Reasons behind not showing real performance	High stress/anxiety	4			
	Inability to feel the exam atmosphere	2			
	Having problems in exam management	1		9	
	Inability to get used to the exam platform	1			
	Having concentration problem	1			
Other comments	Occasionally though not always	4			
	Not having face-to-face exam experience	2		7	
	Change according to situation/psychology	1			

In Table 4.10, the students' views on whether they can show their real performance in online exams were presented. As seen in the table, while most of the students (59,43 %) believed to show their real performance in online exams, a part of students (29,25 %) were not of the same opinion. When the statements of students who believed to show their real performance were examined, it was seen that they claimed to show equal performance with face-to-face exams. Those students noted that stress-free exam procedure was a significant factor for showing their real performance in online exams. Moreover, some students indicated that they showed performance even better than face-to-face exams. Referring that they showed their performance in online exams, some students stated that the difference was not in the exams, but in the way the exam was administered; therefore, their performance was not affected. In the same vein as in this statement, some students highlighted that performance was knowledge-based, not platform base.

The students put forth five reasons for not showing their real performance in online exams. Among these reasons, the first and most expressed one was high stress/anxiety. According to them, since online exams caused high stress/anxiety, they could not show their actual performance in the exams. Some students stated that the lack of exam atmosphere in online exams was another reason which caused to decrease their performance. In addition to these reasons, they also stated that having problems in exam management, inability to get used to the exam platform and having concentration problems were other reasons they raised why they couldn't show their real performance in online exams.

Some of the students' statements on this theme were as follows:

I think that my online and face-to-face exam performances are the same. (SM8, Showing equal performance with face-to-face exams)

Yes, I think so. I think I will get almost the same results in face-to-face exams. (SF42, Showing equal performance with face-to-face exams)

Yes, I think because there is no stress and anxiety. (SM24, Stress/anxiety-free exam procedure)

Yes, because I am not stressed compared to the face-to-face exam. (SF123, Stress/anxiety-free exam procedure)

Yes. It even gets better. (SF7, Showing better performance than face-to-face exams)

Yes, I think I can show my real performance in online exams. (SF95, Showing better performance than face-to-face exams)

Yes, I think because I give the same effort in the online exam as in the face-to-face exam and the exams do not change. (SM116, The difference is not in the exams, but in the way the exam is administered.)

I don't think there is a difference. Exam is exam. (SF118, The difference is not in the exams, but in the way the exam is administered.)

Yes. Online or face-to-face, even if the exam platforms change, the answers to the questions are still within the knowledge of the student. (SF1, Performance is knowledge-based, not platform-based)

Yes, I think so. Because the exams already want the information provided us beforehand, not anything extra. (SF83, Performance is knowledge-based, not platform-based)

I don't think because I'm more stressed than when I'm in the face-to-face exam. (SF32, High stress/anxiety)

No, because I get too stressed. (SM73, High stress/anxiety)

No, because it's not the same atmosphere. (SM73, Inability to feel the exam atmosphere)

No, because if something goes wrong I get anxious trying to fix it up and it's negatively impacting my exam. (SF99, Having problem in exam management)

No, for example, I am not used to reading from the screen, there is a desire to underline sentences. (SF108, Inability to get used to the exam platform)

No, I'm having trouble focusing. (SM43, Having concentration problem)

From time to time. (SF91, Occasionally though not always)

Since I did not perform in the face-to-face exam, I cannot say for sure. (SF72, Not having face-to-face exam experience)

I don't know, it depends on the current psychology. (SM74, Change according to situation/psychology)

Table 4. 11. *Students' views on the skill/skills they have the most difficulty in online exams*

Skill/skills that students have the most difficulty with in online exams	Number of participants	Total number of participants
Listening	52	
Speaking	10	
Writing	9	107
Reading	6	

In Table 4.11, students' views on the skill/skills they had the most difficulty in online exams were presented. Accordingly, the skill that students had the most difficulty in the online exams was listening (*f*=52) which was respectively followed by speaking (*f*=10), writing (*f*=9)

and reading (*f*=6). When the students' statements on this theme were examined, it was revealed that listening was the hardest section in online exams since it could easily be affected by internet connections. The students also underlined this issue in their statements and noted that they had difficulty in understanding listening audios due to poor internet connection and insufficient technical infrastructure. In speaking, the students pointed out that they had an extreme level of anxiety, so they couldn't gather their words in the sessions. Similar to speaking, they also had stress in writing and couldn't organise their work. In reading, on the other hand, they underlined the difficulty of maintaining reading from the screen. Besides, they complained about distracting sounds/noises of their colleagues.

Some of the students' statements on this theme were as follows:

The listening part is because sometimes the sound can go away or the time may be short to write the answers. (SF2, Listening)

Listening because sometimes there are audio interruptions and therefore we can miss it from time to time. (SF20, Listening)

Listening, I definitely had trouble with issues such as connection, sound level in every exam. (SM49, Listening)

Speaking. Because I couldn't gather my words from excitement in 1 minute of thinking. (SF42, Speaking)

Speaking because it is much more stressful than other exams. (SF95, Speaking)

Writing exam because I can't be comfortable writing something, it's like I'm implying that I'm cheating. (SM8, Writing)

Writing. Because I don't have an idea right away. While writing the introduction, I can't start easily. Maybe 10 minutes more time would be great. (SF128, Writing)

Reading. It's hard for me to read from the screen. (SF93, Reading)

It may be reading because sometimes there are distracting noises. (SF118, Reading)

Table 4. 12. *Students' views on whether the problems they encounter in online exams are due to the administration way of exams or general problems*

Are the problems you encounter in online exams due to the online exam or are they general problems?		Online		General	
		f	%	f	%
		67	62,62	30	28,04
			Number of participants	Total number of participants	
Due to online examination	Connection problems/technical issues		26		
	Audio/noise problem		5		
	Disconnection stress/fear		2	23	
	Location/place inconsistency		1		
	Lack of classroom environment		1		
General problems	Stress/Anxiety		9		
	Time management		3	9	
	Sense of inadequacy		1		

In Table 4.12, students' views on whether the problems they encountered in online exams were due to the administration way of exams or general problems were presented. According to the results, while most of the students (62, 62 %) stated that the problems they faced in online exams occurred due to the fact that they were held online, the rest of the students (28, 04 %) noted that they were general problems. When the students' statements on this them were examined, it was revealed that since the exams were held online, they faced problems in five categories: Connection problems/technical issues, audio/noise problems, disconnection stress/fear, location/place inconsistency and lack of classroom environment. The students who stated that the problems they faced were general problems put forth problems in three categories: Stress/anxiety, time management and sense of inadequacy.

Some of the students' statements on this theme were as follows:

I had internet problems once, which is a problem I only have with online exams. (SM13, Connection problems/technical issues)

Technical problems occur because the exam is online but the stress part is about me. (SF95, Connection problems/technical issues)

The sound problem does not attract much attention in the classroom environment, but the necessity of turning on the sound on the computer is very disturbing because there is always a sound coming from the computer and it draws attention. Everyone can register individually in a separate class. (SF34, Audio/noise problem)

There is no problem other than the listening part. When there are friends asking questions during the exam, it causes us to hear and lose concentration because it is online. In the face-to-face exam situation, quieter communication with the teacher during the exam might have been possible. (SM48, Audio/noise problem)

The only problem is, will the internet go out? Is the electricity cut off? I think that the school should make its students more comfortable in this regard. (SF55, Disconnection stress/fear)

There was a constant fear that the internet would break. (SM60, Disconnection stress/fear)

The place we take exams sometimes might be inconvenient. (SM74, Location/place inconsistency)

As I mentioned in the previous questions, we are alone and I think there is too much stress. (SF84, Lack of classroom environment)

I probably would have the same anxiety in face-to-face exam as I did in the speaking. (SF42, Stress/Anxiety)

Anxiety is something that happens in general. (SF97, Stress/Anxiety)

Generally, I have a time management problem. (SF2, Time management)

I'm not very good at listening. (SF65, Sense of inadequacy)

Table 4.13. *Students' preferences for exam types (online or face-to-face)*

		ONLINE		FACE-TO-FACE	
If you had a chance, which one would you choose: online or face-to-face exams?		*f*	**%**	*f*	**%**
		58	55,24	47	44,76

		Number of participants	Total number of participants
Online	Less stressful exam atmosphere	24	47
	Pandemic conditions	12	
	Flexibility of place	4	
	Familiarity	4	
	Equivalency	2	
	More organised and disciplined exam conditions	1	
Face-to-face	Technical/connection issues-free exam procedure	8	30
	More comfortable/ less stressful exam atmosphere	4	
	Better concentration	4	
	A more efficient type of exam	3	
	Providing real exam atmosphere	3	
	Familiarity	3	
	Environmental factors	3	
	Ability to show real performance	1	
	Hassle-free exam process	1	

In Table 4.13, students' preferences for exam types were presented. As seen in the table, most of the students preferred online exams (55, 24%) to face-to-face exams (44, 76%). When the students' statements were examined, it was revealed that the students preferred online exams by citing six features of these exams. Among these features, the most expressed feature was "less stressful exam atmosphere". This feature was respectively followed by pandemic conditions, the flexibility of place, familiarity, equivalency and more organised-disciplined exam conditions. The students who prefer face-to-face exams, on the other hand, presented nine reasons for their choices. Among these reasons, the technical/connection issues-free exam procedure was the most expressed one. It was respectively followed by a more comfortable/ less stressful exam atmosphere, better concentration, a more efficient type of exam, providing real exam atmosphere, familiarity, environmental factors, ability to show real performance and hassle-free exam process.

Some of the students' statements on this theme were as follows:

Online exams

I would prefer online exams because this is how I can control my stress level. I get a lot more stressed when face-to-face, so I am satisfied with this training. (SF21, Less stressful exam atmosphere)

There are no situations such as being late for online exams. In addition, we can take the stress-free exam by entering our home, as we feel psychologically more comfortable and in a protected area. (SM54, Less stressful exam atmosphere)

I would prefer online exams for current conditions. (SF33, Pandemic conditions)

I would prefer online this year because there is a pandemic and I live in a different city. (SF102, Pandemic conditions)

I would prefer online exams, I think it is much more comfortable to take the exam in my own home. I have a special health condition, so I never get tired. (SF50, Flexibility of place)

I can say that online exams are a little more advantageous because we are more comfortable while preparing, we do not spend our time on transportation. (SM131, Flexibility of place)

Online. Because we started online, it can be difficult to get used to face-to-face. (SM35, Familiarity)

Online because we didn't take our training face-to-face, and I would be stunned if I took the exam face-to-face. (SF40, Familiarity)

I would prefer online exams because I think online exams are more organized and disciplined. (SF83, More regular and disciplined exam conditions)

Face-to-face exams

Face to face because online exams can have internet problems and our actions are prone to misunderstanding. (SF118, Technical/connection issues)

I prefer face-to-face exams because I do not encounter any problems I have online, I do not experience excessive stress and I can fully focus on the exam. (SF130, Technical/connection issues)

Face to face. Better concentration, less stress. (SF38, More comfortable/ less stressful exam atmosphere)

Face to face. Less anxiety, at least I am not affected by events against my will, such as internet disconnection. (SM60, More comfortable/ less stressful exam atmosphere)

Face-to-face exams. It is more comfortable to take an exam just by focusing without cheating. (SF5, Better concentration)

I prefer face to face, I adapt to the exam with less fatigue. (SF124, Better concentration)

Face to face. It will be more efficient. (SF108, A more efficient type of exam)

I would prefer face-to-face exams. I am a competitive and ambitious person, so I like to pause from time to time in face-to-face exams and watch how my competitors fail the exam in a sweat. I also think that face-to-face exams are more accurate for general students. Face-to-face exams are more concrete and useful. (SM133, A more efficient type of exam)

I prefer face-to-face exams, it provides the exam atmosphere better. (SF19, Providing real exam atmosphere)

Face to face. I think face-to-face exams are more effective for motivation and a general feeling of the school. (SM56, Providing real exam atmosphere)

Face to face is best because it is a situation we are used to. (SF17, Familiarity)

Face to face because it's a system I'm used to. (SF31, Familiarity)

I would prefer it to be face to face. Because at home, my family can break into the room all the time, or they can make noise from the other room, or it can be the sound of renovations. (SF20, Environmental factors)

Definitely face to face because I know I can show my performance. (SF84, Ability to show real performance)

I would prefer face-to-face exams because I think they are more seamless. (SM12, Hassle-free exam process)

Table 4.14. *Students' views on the skill/skills they wish to continue with online exams*

Skills/skills participants would like to continue with the online exam	Number of participants	Total number of participants
Speaking	47	
Writing	37	
Reading	32	107
All four skills	29	
Listening	10	
None	3	

In Table 4.14, students' views on the skill/skills they wished to continue with online exams were presented. According to the findings, most of the students (*f*=47) stated they would like to continue testing their speaking skill online. It was followed by writing (*f*=37), reading (*f*=32), all four skills (*f*=29), listening (*f*=10) and none of the skills (*f*=3). When the students' statements on their choices were examined, it was found that they clarified their choices by showing the problems in listening. Since listening was affected more by internet connection and technical issues than other skills, they underlined this case in their statements. Those who did not want to continue online reading exams pointed out the difficulty of reading passages from the screen and distracting effect of sounds/noises from other students.

Some of the students' statements on this theme were as follows:

Writing because it requires good focus and I focus better on the online exam. (SF7)

Speaking, writing and listening because we don't have to look at the screen for a long time. It is difficult to read the paragraph on the phone screen while reading. (SF40)

Reading, speaking and writing. Because sometimes we may encounter problems in listening. (SM54)

Reading and writing because in others, internet and technological difficulties may arise. (SM62)

I want all of them except listening. I can understand very well. (SF102)

I would prefer reading because it is a skill that one can develop oneself. (SF84)

I guess it would be writing because I'm at home so I can write and erase more comfortably. (SF96)

I wish the reading and speaking exams were online. The writing exam is the same or even more tiring than face-to-face training. Apart from this, many problems may occur in listening, but reading and speaking can be done online in a more comfortable and safer way. (SM133)

4.3. Findings for the Third Research Question

In this part, the findings related to the second research question "What are the instructors' views on online testing practices in foreign language education?" were presented. In the related research question, the instructors' views on online exams were examined from various aspects. Accordingly, the instructors' views on the practicality of online testing practice in their institutions (Figure 4.1), the problems the instructors faced during the online testing practices (Figure 4.2), the advantages and disadvantages of the online testing practices (Figure 4.3) and the alternative models/techniques proposed by the instructors to use in online testing practices (Figure 4.4) were examined and the findings were presented together with some selected statements of the instructors.

Practicality of online testing practices

Practical (13)
They prevent cheating.
Open-ended questions are used for reliability.
They are sufficient under the current conditions.

Not practical (11)
It is difficult to grade.
They are not utilised from online testing systems.
They are adapted versions of face-to-face exams.
It is difficult to proctor.
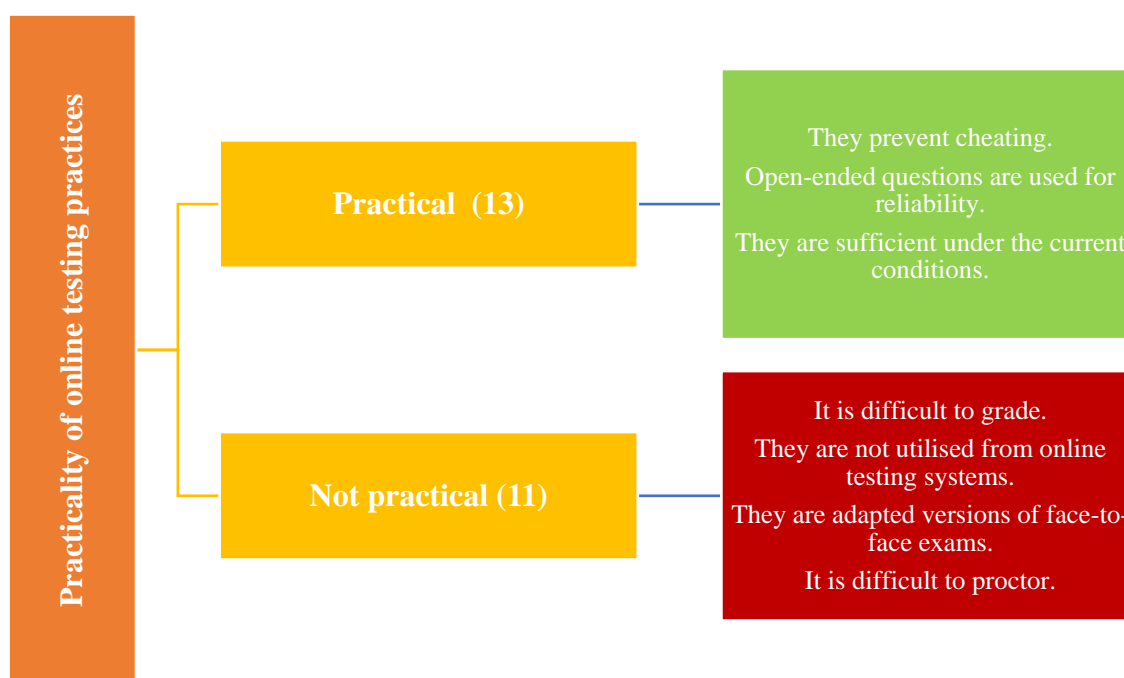
Figure 4. 1. *Instructors' views of the practicality of online testing practice in their institutions*

In Figure 4.1, instructors' views on the practicality of online testing practices in their institutions were given. Accordingly, 13 instructors believed that online testing practices in their institution were practical since they prevented cheating, open-ended questions were used for reliability and they were sufficient under the current conditions. The instructors indicated that they prevented cheating, citing the use of open-ended questions in online exams in their institutions. For some instructors, using open-ended questions was also an indicator of reliability. Besides, one of the instructors underlined that their online testing practices were

adequate under the current circumstances but updates should be done in order to advance the quality of exams.

On the other hand, 11 instructors thought that they were not practical since it was difficult to grade, the exams were not utilised from professional online testing systems, they were adapted versions of face-to-face exams and it was difficult to proctor. The instructors pointed out the difficulty of grading by showing the use of open-ended questions and double-checking process. For some instructors, their institutions' online exams were not practical because they did not get professional support in terms of the testing platform. The instructors also stated that online exams of their institutions were not specially prepared for this purpose and that these exams were adapted versions of exam samples used in face-to-face exams. According to the instructors, online testing practices in their institution were not practical because there were many things to do while proctoring and that was quite hard to manage. Some of the instructors' statements on this theme were as follows:

Different ways have been tried to find the best way since the beginning of online education, and I think it's practical because we try to prevent cheating issues. (IF23)

I think they are practical since the open-ended question format is preferred to evaluate students' receptive skills. (IF12)

The existing practices are practical enough for now, however change is a must in education. There is always room for development. Even if the existing practices help a lot, we should seek for more and new practices. (IF11)

They are at the point of taking exams, and not at the point of checking. The checking process takes much time and using paint for checking is not practical. (IF17)

I think that they are far from practical as there is no use of online testing systems, we send papers to the students via email or teams, which cannot be counted as online testing practice. (IF9)

Not practical. Adapted versions of the usual exams. (IF13)

I find them a bit hard for teachers to proctor and grade. (IF14)
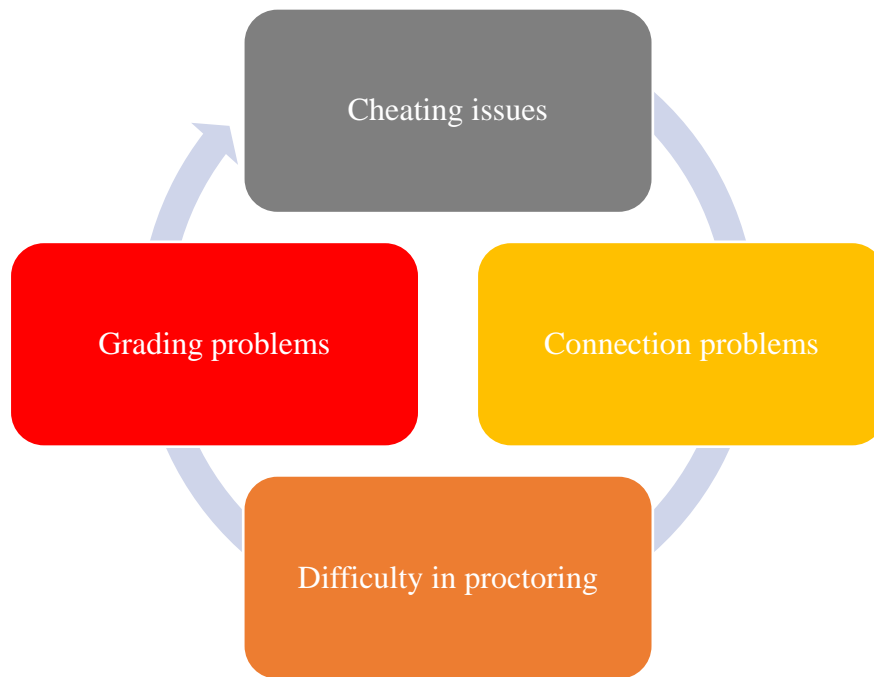
Figure 4. 2. *Problems the instructors faced during the online testing practices*

In Figure 4.2, problems the instructors faced during the online testing practices were presented. As a result of analysing the instructors 'statements on the theme, the problems were grouped under five categories: Cheating issues, connection problems, difficulty in proctoring, limited angle of screen and grading problems. In the first category, the instructors complained about the cheating issues occurred during the exams. They mainly believed that it was quite hard to detect cheating cases and prevent students from this action. The second problem they faced was connection problems. According to the instructors, connection problems were one of the greatest drawbacks of online exams since they unpredictably occur while testing and negatively affects students' grades, especially in listening tasks. Another problem they faced was difficulty in proctoring. In this category, they underlined the processes they followed before and after the exams. They also noted that since they had a limited angle of the screen, they stated that they could not fully control the students and had difficulty in making judgments about cheating etc. The last problem expressed by the instructors was the grading problem which occurred since open-ended questions were used in exams.

Some of the instructors' statements on this theme were as follows:

They can cheat during writing exams and in listening parts sometimes we have problems because of connection or not hearing well. (IF10, Cheating issues)

For me, the biggest problem is cheating. You never know how students cheat and even if we have some strict rules about the exam procedure, students find a way to cheat. Maybe we can ask students to use two different devices one is for reading the questions and the other for camera purposes but the camera

should show the first device screen and the student with answer paper so we can understand if they are cheating or not. Still, it's too complicated. (IF11, Cheating issues)

Cheating attempts are not easy to detect. If students have advanced computer skills, it is very easy to evade being caught. Also, when students claim that they have technical problems, we do not have a way to tell whether they are telling the truth. (IM21, Cheating issues)

Internet connection problems, especially problematic during listening exams. (IF4, Connection problems)

Bad internet connection problems especially on rainy days, dark rooms or unclear, blurry images, having students who have only one device during the exam and not being able to see the screen that students are controlling. (IF11, Connection problems)

Seeing all the students in a class at the same time is impossible. Internet connection problems might occur. (IF1, Difficulty in proctoring)

Monitoring students is way more challenging since there is no physical environment and eye contact. (IM26, Difficulty in proctoring)

Questions are open-ended but it's hard to grade them. (IF24, Grading problems)
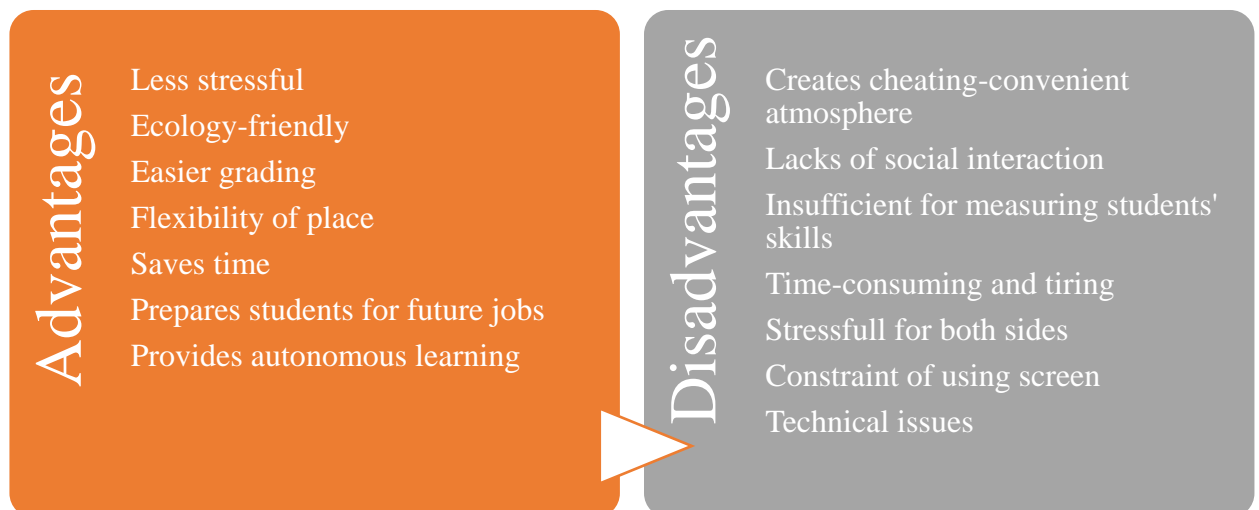
Advantages
- Less stressful
- Ecology-friendly
- Easier grading
- Flexibility of place
- Saves time
- Prepares students for future jobs
- Provides autonomous learning

Disadvantages
- Creates cheating-convenient atmosphere
- Lacks of social interaction
- Insufficient for measuring students' skills
- Time-consuming and tiring
- Stressfull for both sides
- Constraint of using screen
- Technical issues

Figure 4. 3. *The instructors' views on the advantages and disadvantages of the online testing practices*

In Figure 4.3, the instructors' views on the advantages and disadvantages of the online testing practices were presented. As seen in the figure, the instructors noted seven advantages and disadvantages for online testing practices. The first advantage expressed by the instructors was that the online exams were less stressful. Another advantage of online exams was that they

were ecology-friendly. Referring that no physical paperwork was not used in online exams, the instructors considered it as a significant aspect of online exams. Since everything was prepared and conducted online, they also thought grading was easier in online exams compared to face-to-face counterparts. Connecting online from anywhere was also regarded as an appealing advantage by the instructors. Besides, they believed that online exams saved time, prepared students for future jobs and provided autonomous learning. When it came to disadvantages, the first and most expressed disadvantage of online exams was they created a cheating-convenient atmosphere. While some instructors stated it was no possible to prevent cheating, others suggested alternative models in order to avoid cheating. Since everything was conducted online, lack of social interaction was another disadvantage of online exams expressed by the instructors. For some instructors, online exams were not sufficient for fully measuring students' skills, so they considered it as a disadvantage. Referring to grading processes, the instructors also noted that online exams were time-consuming and tiring. While online exams were considered as less stressful while expressing advantages, they were regarded as stressful for both students and instructors in disadvantages. This finding showed the difference of individual perspectives. On this theme, the constraint of using screen and technical issues were other disadvantages expressed by the instructors.

Some of the instructors' statements on this theme were as follows:

Advantages: students are less stressed, teachers are more relaxed, too. (IF4, Less stressful)

Advantages: students are less excited, no physical paperwork. (IF17, Less stressful)

The best advantage of online testing is that it is ecological as we don't waste thousands of paper. (IF11, Ecology-friendly)

Pros: it's safe for students' and teachers' health during Covid and also environmentally friendly as they don't have to write on papers and we don't print exams. (IF20, Ecology-friendly)

Since everything is done online it's easier to calculate and get the results. (IF3, Easier grading)

Less time consuming as computers make the assessment. (IF22, Easier grading)

It prepares students for future jobs, it can be convenient regardless of the actual location. (IM2, Flexibility of place, Prepares students for future jobs)

I don't see any advantages except being able to test students from different locations. (IF14, Flexibility of place)

It saves time and nature. (IF1, Saves time)

Online examination system gives students the ability to take the responsibility for their own learning and to see their learning process because they can keep their exam paper even after the exams. (IF12, Provides autonomous learning)

However, cheating and lack of social interaction are considerable issues. (IM2, Creates cheating-convenient atmosphere, Lacks of social interaction)

It's not possible to stop cheating. (IF7, Creates cheating-convenient atmosphere)

It is not enough to measure students' skills. (IF10, Insufficient for measuring students' skills)

Disadvantages are that it's time-consuming, requires lots of unnecessary effort and is unsuitable for modern times. (IF9, Time-consuming and tiring)

Stressful for students as more hiccups they can't control. Stressful for instructors too. (IF8, Stressful for both sides)

The constraint of using the screen for everything. (IF6, Constraint of using screen)

One drawback of an online examination system is that both instructors and students might have some challenges in technology, network connection speed etc. during the exams. (IF12, Technical issues)

Alternative models/techniques for online testing

- Assigning projects/portfolios rather than tests
- Creating an online test like TOEFL iBT
- Assessing speaking skills in groups instead of one by one
- Grading based on course works
- Integrating different types of questions
- Blocking students' devices during the examination
- Using plagiarism detection programs
- Having a reliable electronic proctoring system
- Presenting questions by shuffling
- Making a testing practice with less students at one time
- Sending mirrors for more effective proctoring

Figure 4. 4. *Alternative models/techniques proposed by the instructors to use in online testing practices*

In Figure 4.4, the instructors' views on alternative models/techniques to use in online testing practices were presented. Accordingly, the instructors suggested twelve models/techniques that they thought would have a better examination process if implemented.

The first suggestion was "assigning projects/portfolios rather than exams". The instructors believed that project/portfolios could be an appealing option to be used in online education instead of exams. Considering the instructors' views on the disadvantages of online exams, they might have suggested projects/portfolios since there was not cheating concern in this type of assessment. The second model/technique proposed by the instructors was creating an online exam like TOEFL iBT. This suggestion showed that some parts of the online exams prepared by their institutions were not sufficient for online exams, therefore they might have felt the need to make such a suggestion. The third suggestion was grading based on course works. This suggestion of the instructors showed similarity with the first suggestion. Either way, it was intended that students would be assessed through the tasks or projects assigned to them, rather than through testing. The following suggestion was integrating different types of questions. Referring to the fact that the questions in the exams merely consisted of open-ended questions, the instructors suggested that different types of questions should be included in the new exams. The suggestion "blocking students' devices during the examination" was significant for preventing cheating attempts. This suggestion showed that they were aware of the students' behaviours, suspected the students of cheating. Referring to the plagiarism issues in writing exams, the instructors also suggested using plagiarism detection software. As it was stated in the previous questions, the instructors noted that they needed a reliable proctoring system since it was a complicated and tiring process for them. Another suggestion of the instructors was to present questions by shuffling which was also a precaution to prevent cheating issues. One of the instructors proposed that organising the exam with fewer students could also be a good way to advance the quality of testing practices. Another instructor suggested using the technique of a private university located in Southern Turkey and added that they could also send mirrors to the students and asked them to use the mirrors during the exams.

Some of the instructors' statements on this theme were as follows:

I would assign students with projects rather than timed exams. (IF1)

Something like TOEFL computer exams where we send students links, they have a timer and only one screen for the exam. Maybe add a camera to make sure the students are answering alone and their session is videotaped. (IF20)

I would change the speaking exam. Instead of testing the students one by one I'd test them in groups of three or at least in pairs. It should be more like a conversation instead of a presentation. (IF5)

For genuine testing progress, more realistic meaningful grades from course work. (IF8)

A system that has different varieties of question types such as multiple-choice, matching etc. which calculates the results itself. We can also ask 1 or 2 open-ended questions I mean the system can have that option as well and teachers can assess only those. (IF9)

Exams can be done once for each level which are prepared, applied and assessed online without using any papers. In a way that blocking the usage of other websites during exams. (IF13)

Even if it takes more time to mark the exam paper, I am glad that we don't use multiple choice questions in our online examination system because exams and quizzes assess low-level learning. For the writing exams; however, plagiarism detection programs can be used for the benefit of students and teachers. (IF12)

Having a reliable electronic proctoring software would help a lot. (IF14)

Cameras must be on at all times and shuffle the questions for each student. (IF22)

I would design a testing practice with fewer students at one time and focus on the production skills more. (IF23)

For the testing part, at Bilkent University for example they sent mirrors to students to have more effective monitoring that is a brilliant idea I believe. (IM26)

CHAPTER V

CONCLUSION, DISCUSSION AND SUGGESTIONS

5.1. Introduction

In this chapter, the results obtained in the light of the findings of the research are presented with the support of the research findings in the relevant literature. After the conclusion and discussion section, suggestions for researchers and practitioners are put forth.

5.2. Conclusion and Discussion

This research is a case study in which instructors and students' views on online teaching practices in foreign language education were sought. As a result of examining the data collected through online opinion forms, the following results were achieved:

Four different aspects were investigated within the first research question. In the first aspect, it was aimed to unearth the relationship between students' views on their success situations in face-to-face exams and their demographic characteristics (gender, age, English level, and computer skills). The analysis results showed that the students' views on success situations in face-to-face exams did not differ statistically based on their gender, English level and computer skills. However, it was found that their views on success situations in face-to-face exams meaningfully changed according to their ages. It was determined that there was an inverse relationship between the age of the students and their views on their success in face-to-face exams, and as their age increased, their views on their success levels decreased.

In the second aspect, the relationship between the students' views on success situations in online exams and their demographic characteristics were investigated. When the results were examined, it was found that while the students' views on success situations in online exams did not significantly change according to their gender and age, it changed in a significant way based on their English level and computer skills. This differentiation merely occurred between the students at pre-intermediate level and those at intermediate, upper-intermediate and advanced level. Among these students, pre-intermediate students had the lowest belief concerning the success situation in online exams. In other words, the higher the English level the students were, the higher their belief in success in the online exam would be. When the results on computer skills were taken into consideration, it was seen that there was a direct relationship between the

students' computer skills and their views on success in online exams. Accordingly, as the students' computer skills increased, their views on success in online exams increased as well. The findings of this study varied from Urgun (2019) and Wallace and Clariana's (2005) studies in which they found female students achieved better scores than males in the computer-based exams.

In the third aspect, the relationship between students' views on their stress level in face-to-face exams and their demographic characteristics (gender, age, English level, computer skills) was scrutinised. The analysis results indicated that female students had higher stress in face-to-face exams than males. In contrast, no relationship was found between the students' other demographic characteristics (age, English level and computer skills) and their stress level in face-to-face exams.

In the fourth aspect, the relationship between students' views on their stress level in online exams and their demographic characteristics (gender, age, English level, computer skills) were examined. Like in face-to-face exams, it was determined that female students had higher stress levels than male students in online exams. Besides, it was seen that the students' computer skills were also a significant variable changing students' stress in online exams. The results attested that there was an inverse relationship between the students' computer skills and their stress levels in online exams. Accordingly, the students who had lower computer skills had higher stress in online exams. The results also showed that the students' stress in online exams did not significantly change in terms of their age and English levels. In a similar vein, Özturan (2016), who made a comparison on online versus paper-based exams, found that the students having higher computer skills had less anxiety in online exams.

Through the second research question, the students' views on online exams were examined through the statements in opinion forms from nine different aspects. In the first aspect, the students' views on the advantages and disadvantages of online exams were sought. As a result of examining their statements through content analysis technique, three advantages and six disadvantages of online exams were identified. Among the advantages, the most expressed advantage by the students was that online exams provided a less stressful exam atmosphere than face-to-face counterparts. Besides, the students also underlined the advantage of the flexibility of place. Although not as much as the other two advantages, the fact that online exams allowed better concentration was also mentioned by the students as another advantage. When it comes to disadvantages, connection/technical issues was the most expressed disadvantage of online exams. Almost one out of every two students stated that they had

problems with this disadvantage. This was seen in the statements of the students. Students explained their opinions about this disadvantage by giving examples such as disconnection from the internet, power cuts, and problems arising from using a computer without new equipment. Another disadvantage expressed by the students was that online exams caused stressful exam procedures. Whereas most of the students stated that online exams enabled them to take the exams in a less stressful atmosphere, a group of students claimed that online exams led them to feel greater stress which mainly resulted from the fear/anxiety of 'what if my internet connection is lost or if the electricity goes out' and the problems in understanding listening audios due to poor internet connection. The students also stated that online exams created a cheating conductive environment which caused injustice between studying and non-studying students. Another disadvantage expressed by the students was adaptation/concentration problems. They noted that they had adaptation/concentration problems in online exams since they were forced to leave their microphones on during the exams and sounds/noises of their classmates, devices and people at home functioned as distracting in online exams. Communication problems and difficulty in finding a convenient place were the other disadvantages expressed by the students.

In the second aspect, it was aimed at identifying problems the students encountered in online exams. As a result of examining students' statements, problems were grouped under five categories; connection/technical problems, sound/noise problems, stress, camera problems, cheating attempts and insufficient exam time. Connection/technical problem was the most mentioned one among the problems expressed by more than 80% of the students. The second general problem was the sound/noise problem which arose from the rule of keeping the microphone on during the exams. This situation was specifically stated in the statements of the students and they emphasized that even if they did not want to disturb others, they were unintentionally forced to do it since they did not have full control of the process. Another general problem the students faced in the online exams was camera problems which mainly emerged either because there was a requirement for opening cameras during the exams or because they were asked to position their cameras properly to get a clear view in the exams. The students complained that since the cameras were open during the exams, they got stressed more and couldn't focus on their papers. Some students pointed out they had difficulty positioning their cameras properly. Cheating attempts and insufficient exam time were the other general problems, each expressed by one student. As it was stated in the disadvantages of online exams, one of the students emphasised cheating problems in online exams under the general

problems though. Another student noted they had difficulty in completing exams so considered insufficiency in exam time as one of the general problems.

In the third aspect, it was aimed at identifying aspects of online exams that needs changing/improving according to the students. According to the results, most of the students believed that there were aspects of online exams that needs changing/ improving. From the statements of the students who thought there were aspects of online exams that needs changing/improving, 18 aspects were determined. Among these, extending the exam period was the most expressed one. The students especially underlined this aspect because they claimed to have difficulty in completing the tasks in the exams and had problems merely because of the administration way of the exams. Therefore, they believed that extending the exam period could be a good compensator to overcome these situations. In addition to this, removing the requirement for microphones/cameras to be on were the following aspects that need changing. Like in the previous questions, the students expressed they had concentration/adaptation problems due to these requirements. Besides, they noted that the requirement for microphones to be on caused noises/sounds arising from other classmates, electronic devices and people in the environment. They added that the requirement for cameras to be on brought about stress among them. Another aspect of online exams in this theme was conducting additional listening which was specifically expressed by the students since they thought listening was the skill most affected by the internet connection. Some students noted they couldn't understand the audio just because they did not have a stable internet connection in their environment, so providing additional listening (three times) could solve this problem for them. Moreover, some students specified that changing question types could be another aspect of online exams. Since the exams prepared by the university board consisted of open-ended questions, some thought online exams should cover a variety of question types to properly assess their skills. Removing the camera angle rule and bending the rules were the other aspects stated by the students who believed that these kinds of rules caused an increase in their stress levels. The students also stated that fixing technical issues and changing the exam platform were the aspects that needed changing. In their statements, they emphasised the technical issues such as internet connection and quality of server and asked the university to advance these issues for maintaining a better exam procedure. By asking to change the exam platform, they recommended the university utilise alternative platforms such as websites. Within the scope of this research question, the students also suggested the following aspects, each expressed by one student: Shortening the exam time, Conducting the speaking exam with several students rather than one-on-one, adding grammar section, not treating as cheating in disconnection, increasing the degree of difficulty of the

questions, increasing the number of cameras, ensuring the instructions were understood, preparing speaking questions of equivalent difficulty and omitting the listening section.

In the fourth aspect, the students were asked whether additional time should be given in online exams. The results showed that the majority of the students thought that additional time should be given in online exams. For this opinion, they provided justifications under 11 categories. The first and most expressed reason was the difference in connection/technical features in which they claimed that everyone didn't have the same quality of internet connection and electronic devices, so additional time should be given in online exams. For students, difficulty in concentrating on the exam was another reason for additional time. Besides, they put forward that they had different environmental conditions and didn't have full control over them so additional time should be given in online exams. Some students specified the areas that needed additional time. For some, it was the listening section due to the understanding problems that aroused from the internet connection. For others, it was the reading section since it was hard to maintain reading tasks from the screen. Moreover, some noted that it was the writing section since nothing came to their minds and they had difficulty in continuing writing without looking at the screen. In addition to these, some students cited the difference in the exam platform to defend their view of why additional time should be given for online exams. Some students in the research group, on the other hand, stated that everyone did not have the same computer skills, so it was necessary to give additional time in online exams. Some students, who focused on a different aspect of the online exams, stated that they were warned frequently during the exams and therefore they lost their concentration and noted that giving additional time would be the right step. As it was mentioned in the disadvantages of online exams, some referred to high stress/excitement online and therefore asked for additional time in online exams. The last reason expressed by the students was exam familiarization and device control in which they emphasized the time wasted.

In the fifth aspect, the students were asked whether they could show their real performance in online exams. When the results were examined, it was found that the majority of the students showed their real performance in online exams. On the contrary, a group of the students claimed not to show their actual performance in online exams. After scrutinizing statements of both groups, five reasons for showing real performance, five reasons behind not showing real performance and three reasons for other issues were determined. The students who claimed to show their real performance in online exams mostly stated that they showed equal performance with face-to-face exams. Some students noted that stress/anxiety-free exam

procedures assisted them to show their real performance in online exams. A group of the students thought their performance in online exams was better than in face-to-face exams. By stating that the difference was not in the exams but in the way the exam was administered, some students referred to the fact that they showed the same performance in online exams. The last reason in this category was the motto "performance is knowledge-based, not platform-based". Through this quotation, the students implied that any change in the exam platform did not affect their performance if they had sufficient knowledge in the relevant subject. When the reasons behind not showing real performance in online exams were examined, it was found that high stress/anxiety was the most expressed reason by the students. The students also gave the following justifications for not showing real performance in online exams: inability to feel the exam atmosphere, having problems in exam management, inability to get used to the exam platform and having concentration problems. Within the scope of this research question, there were three reasons for other issues. Among these, the students most expressed that they occasionally, but not always, showed their real performance in online exams. Some students clarified that they didn't have face-to-face exam experience at the university level, so they couldn't make a comparison between them. One of the students, on the other hand, stated that his performance changed according to situation/psychology. Similar results were revealed in many studies (Anakwe; 2008; Bayazıt; 2007; Campton, 2004; Candrlic, Ktic & Dlab, 2014; Jeong, 2014; Still & Still 2014) in the literature. As stated in this study, there was no significant difference between students' online and face-to-face exam performances in these studies. In other words, these studies attested that the students showed equal performance in both exams.

In the sixth aspect, it was targeted to explore the skill/skills the students had the most difficulty in online exams. The results revealed that listening had been the most challenging skill for students in online exams. Since it was quite sensitive to the internet connection, the students expressed they had difficulty in following listening records and sometimes needed additional listening due to muffled sounds during the listening section. Listening was followed by speaking, writing and reading respectively. Although they didn't give a clear reason for speaking related to the online examination, they voted for the second since they had difficulty in organising the sentences and got stressed during the section. Similarly, the students noted that nothing came to their minds during the writing section and it was hard for them to complete writing tasks without looking at the screen which was a rule set by the university to avoid cheating and plagiarism. For reading skills, the students underlined the difficulty of maintaining reading tasks from the screen.

In the seventh aspect, the purpose was to clarify whether the problems students encountered online exams were due to the administration way of exams or general problems. According to most of the students 'views (62, 62 %), it was because the exams were administered online. The rest of the students (28, 04 %) thought that they were general problems. From the statements of the students who thought the problems were due to the online examination, five categories were determined. Among these, connection problems/technical issues were the first and most expressed by the students. It was followed by audio/noise problems, disconnection stress/fear, location/place inconsistency and lack of classroom environment. When the statements of other students were checked, three general problems were determined. Stress/anxiety was the most common problem the students faced in the exams. It was followed by time management and a sense of inadequacy.

In the eighth aspect, the students' preferences for exam types (online or face-to-face) were examined. The results showed that most of the students (55, 24 %) preferred online exams to their face-to-face counterparts. The rest of the students (44, 76 %) wished to continue with face-to-face exams. The students choosing online exams put forth six reasons for their choices. Among these reasons, the fact that online exams created a less stressful exam atmosphere was the most expressed one by the students. The students also added that they would choose online exams because there were pandemic conditions, online exams provided the flexibility of place, they got used to taking exams online from the beginning of their university education, they thought online-or face-to-face they were equivalent and online exams provided more organised and disciplined exam conditions. When the statements of the students choosing face-to-face exams were examined, nine reasons were determined. Out of these reasons, technical /connection issues had the greatest repetition value. Some students stated that in face-to-face exams, they had a more comfortable/less stressful exam atmosphere, had better concentration, could show real performance, had a hassle-free exam process and were not affected by environmental factors. Besides, they thought that face-to-face exams were a more efficient type of exams and provided a real exam atmosphere. Like the students choosing online exams, some students in face-to-face exams also noted that they were used to taking the exam face-to-face throughout their educational lives, so they had familiarity with face-to-face exams. Similar to the findings in this study, Çörekçioğlu (2017) stated that the features such as time-saving, environmental friendliness, providing validity and motivation of online exams were welcomed by both teachers and students. On the other hand, he underlined that connection and computer problems were frequently mentioned by both groups.

In the ninth aspect, the students' views on the skill/skills they wished to continue with online exams were investigated. The analysis results showed that speaking (f=47) was the most desired skill the students wished to continue with online exams. It was followed by writing (f=37), reading (f=3), all four skills (f=29), listening (f=10) and none of the skills (f=3).

Through the third research question, the instructors' views on online exams were examined from four different aspects. In the first aspect, it was tried to explore their views on the practicality of online testing practices in their institutions. According to the results, while most of the instructors (f=13) believed that online testing practices in their institutions were practical, other instructors (f=11) thought there were not practical as much as expected. From the statements of the instructors thinking of online testing practices of their institutions as practical, three factors were determined. Firstly, the instructors thought they were practical because they prevented cheating. Besides, open-ended questions were used in the exams which was a good indicator of reliability. Some instructors noted that although they were sufficient/practical under the current conditions, there might be changes/updates for a better assessment. The instructors thinking of them as not practical, supported their ideas with the statements such as the difficulty of grading, not utilising online testing systems, being adapted versions of face-to-face exams and difficulty of proctoring. Due to these reasons, they thought the online testing practices of their institutions were not practical as much as desired.

In the second aspect, it was aimed at identifying problems the instructors faced during the online testing practices. As a result of examining the instructors' statements, the problems faced by the instructors were grouped under four headings: Cheating issues, connection problems, difficulty in proctoring and grading problems. As expected, cheating issues were the most expressed problems by the instructors. For some, it was not possible to prevent these cases since there were students with advanced computer skills or they had limited control over the process. The second one "Connection problems" was the problem that was declared by both students and instructors. In the difficulty of proctoring, the instructors complained about the difficulty and tiring side of the process. In grading problems, the instructors referred to the fact that the exam questions were open-ended, and that the grading process lasted longer than usual and the questions did not have a single correct answer. In addition, they stated that double-checking made the process more difficult. Although the instructors mainly complained about grading and proctoring difficulties, the related studies showed that there were studies (Çelik; 2006, Emir; 2006), İçten; 2006, Jia & He, 2021; Jung & Yeam, 2009; Yağcı; 2012; Yılmaz İnce, 2016) on developing such systems to conduct an effective examination. In particular, the

study of Jia and He (2021) developed an important solution for the cheating problem, which the instructors expressed as the problem they experienced in online exams. Thanks to this artificial intelligence supported system, the students' facial movements and behaviours were examined, and cheating attempts were detected during the exam and deterrent measures were taken for students. On the other hand, online exam software developed by researchers such as Çelik (2006), Emir (2006), İçten (2006) and Yağcı (2012) provided solutions for grading problems expressed by instructors. The online automatic Turkish essay scoring system developed by Yılmaz İnce (2016) took place as an important step in grading the tasks given within the scope of writing skills. If the relevant system is applicable for essays written in a foreign language, perhaps the essay reading workforce, where instructors spend the most time, will be reduced. Hylton, Levy, and Laurie (2016) focused on a different point regarding proctoring. In their study, they found that the unproctored group got higher grades and had the opportunity to be involved in more unethical events. In a similar study, Vazquez, Chiang & Sarmiento-Barbieri (2021) also found that the unproctored group got higher grades and the difference between the groups was greater in face-to-face examinations. As a justification for this, it was shown that the students established a more organized and wider network in face-to-face exams.

In the third aspect, it was aimed at exploring the advantages and disadvantages of online testing practices through the eyes of the instructors. In terms of advantages, seven advantages were declared by the instructors. Among these, the most expressed one was that online exams were less stressful. The instructors also added they were eco-friendly because no paper was used in the exams. Since some parts of the exams were automatically scored by the computers, some instructors underlined that grading was easier in online exams. Another advantage of online exams was the flexibility of place. According to the instructors, online exams also saved their time, prepared students for future jobs and provided autonomous learning. When it came to disadvantages, the instructors firstly complained that online exams created a cheating-convenient atmosphere. They also added that they lacked social interaction, were insufficient for measuring students' skills, were time-consuming and tiring, and were stressful for both sides (students and instructors). Some instructors pointed out that there were technical problems and constraints of using the screen in online exams and that were other drawbacks of these exams. Like the instructors, Çörekçioğlu (2017) reported that teachers and students in his study frequently complained about internet connection and computer problems which were considered as the greatest handicaps of online exams. In that study, similar advantages were

also put forth. As it was found in this study, the participants of Çörekçioğlu (2017) noted such advantages as being time-saving and eco-friendly.

In the fourth aspect, the purpose was to benefit from the experiences of the instructors in online exams. Therefore, they were asked to propose alternative models/techniques to be used in future online testing practices. From the statements of the instructors, twelve alternative models/techniques were revealed. The first one was to assign projects/portfolios rather than exams. The instructors suggesting this model stated that cheating could merely be prevented through this model and there wouldn't be technical issues. Some instructors suggested creating an online exam like TOEFL iBT. As an alternative model, some instructors advised assessing speaking skills in groups rather than one by one. This situation was also declared by the students who stated that testing speaking one by one caused stress and that there was injustice among the difficulties of the questions students faced. Similar to the first alternative model, some instructors noted that grading should be based on course works. While the purpose of the first model was to avoid cheating, the primary goal of this model was to reduce the workload. Since the university mainly used open-ended questions in exams to decrease cheating cases, the students faced with a single type of question, so some instructors recommended integrating different types of questions in future exams. Cheating was one of the first problems the instructors encountered in online exams. Several instructors advised using a system in which students' devices were blocked for other pages/browsers during the examination. In a similar vein, the instructors suggested using plagiarism detection programs especially to check papers in the writing section. Besides, referring to the time-consuming and tiring aspect of proctoring, some instructors stated that a reliable electronic proctoring system should be used in future exams. Related to the cheating issues, some instructors suggested presenting questions by shuffling, others advised the university to send mirrors for more effective proctoring. The last alternative model/technique was to make a testing practice with fewer students at one time.

5.3. Suggestions

According to the results obtained as a result of the research, the following suggestions were put forth for practitioners, policymakers and researchers.

- ✓ Since this research is a case study, it was conducted with a smaller study group compared to quantitative research designs. In order to increase the generalizability of the findings and to reach wider masses, it is recommended to carry out studies that will include quantitative research designs on this or a similar subject.

- Perhaps one of the biggest handicaps of online exams is internet connections/technical issues. Although what can be done to solve these problems is limited, universities can improve the quality of the servers they use in online exams and provide a partial solution to this problem by using systems/software suitable for online exams. In addition, data usage can be minimized so that students connecting from regions with low connectivity can take the exam comfortably and smoothly.

- Another common problem with online exams is cheating. In order to overcome this problem, information exchange can be prevented by using software systems that limit students' computer use during the exam, as suggested by the instructors within the scope of the research. In addition, if the scope of the course is appropriate, techniques such as project and portfolio evaluation, which are among the contemporary evaluation techniques, can be used.

- Considering that students have a less stressful exam process, apart from technical problems/internet problems, online exams seem to have a better process in showing students' real performance. In the field of online exams, which is one of the reflections of the increasing and developing technological developments on education, it will be a good gain for both students and educators to carry out infrastructure development studies and to bring domestic and national infrastructure systems to the education community.

- In order to test the listening skill, which is the skill that students have the most difficulty with in online exams, listening audios can be sent to the students before the relevant section, and a certain time can be allocated for the students to download these files and do listening activities. Students can answer the questions in the relevant section until the countdown process is completed, and when the countdown is over, the answer process can be closed automatically. In this way, students can get rid of internet-related problems in the listening section.

- Online exams are considered as exams that enter our lives quickly and compulsory with the effect of the pandemic. Although we have started a fast process, it is obvious that we will carry out a more intertwined process in the future with the effect of developing technological applications. It is essential to carry out some studies in order not to experience the disadvantages of online exams again. As seen in this study, there is an inverse relationship between students' computer skills and their stress and success levels in online exams. From this point of view, personal development courses can be started to develop students' computer skills, and pre-service and in-service training can be given

to pre-service teachers/ teachers so that they can provide a more effective examination control.

REFERENCES

Abbott, G. J. Greenwood, J. (1985). *The teaching of English as an international language*, Collins: Glasgow and London.

Afacan-Adanır, G., İsmailova, R., Omuraliev, A., & Muhametjanova, G. (2020). Learners' perceptions of online exams: A comparative study in Turkey and Kyrgyzstan. *International Review of Research in Open and Distributed Learning*, *21* (3), 1-17.

Akar-Vural, R., & Somers, J. W. (2012). *Hümanist ilköğretim programları için ilköğretimde drama: Kuram ve uygulama* (2. Baskı). Ankara: Pegem Akademi.

Aksan, D. (2007). *Şiir dili ve Türk şiir dili*. Ankara: Bilgi Yayınevi.

Aksoy, H. (2018). *Uzaktan eğitimde merkezi sınav sistemi* (Yayımlanmamış yüksek lisans tezi). Kırıkkale Üniversitesi, Fen Bilimleri Enstitüsü, Kırıkkale.

Alderson, J., & Banerjee, J. (2001). Language testing and assessment (Part I). *Language Teaching*, *34* (4), 213-236.

Alıcı, D. (2014). Öğrenci performansının değerlendirilmesinde kullanılan diğer ölçme araç ve yöntemleri. S. Tekindal (Ed.), *Eğitimde ölçme ve değerlendirme* içinde (2. Baskı, s. 131–174). Ankara: Pegem Akademi.

Anakwe, B. (2008). Comparison of student performance in paper-based versus computer-based testing. *Journal of education for business*, *84* (1), 13-17.

Angelo, T., & Cross, K. P. (1993). *Classroom Assessment Techniques*. San Francisco: Jossey - Bass.

Arıkan, S., Çelen, Ü., Gülleroğlu, H. D., Gültekin S., Kilmen, S. ve Köse İ.A. (2014). Öğrenme, öğretim ve değerlendirme arasındaki ilişkiler. (Ed: Demirtaşlı). *Eğitimde Ölçme ve Değerlendirme* (s.12-13). Ankara: Tarcan Matbaacılık

Atılgan, H. (2012). *Eğitim bilimleri ölçme ve değerlendirme*. Ankara: Yediiklim.

Atılgan, H., Kan, A. ve Doğan, N. (2017). *Eğitimde ölçme ve değerlendirme*. Ankara: Anı Yayıncılık.

Aydoğdu, M. ve Kesercioğlu T. (2005). *İlköğretimde fen ve teknoloji öğretimi*. Ankara: Anı Yayıncılık.

Bachman, L. F. (2000). Foreword to G. Ekbataniand H. DeLeonPierson, *Learner–Directed Assessment in ESL.* Hillsdale, NJ: Erllbaum.

Bachman, L. F., & Palmer, A. S. (1996), *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.

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

Bahar, M., Nartgün, Z., Durmuş, S. ve Bıçak, B. (2014). *Geleneksel-tamamlayıcı ölçme ve değerlendirme teknikleri öğretmen el kitabı* (6. baskı.). Ankara: Pegem Akademi.

Bailey, K., & Brown, J. (1999). *Learning about language assessment: Dilemmas, decisions, and directions & new ways of classroom assessment*. Learning, 4 (2), 1-8.

Baird, C., Healy, T., Johnson, K., Bogie, A., Dankert, E. W. & Scharenbroch, C. (2013). A Comparison of Risk Assessment Instruments in Juvenile Justice. Madison, WI: National Council on Crime and Delinquency

Bassey, M. (1999). *Case study research in educational settings*. Philadelphia: Open University Press.

Başkömürcü, G. ve Öztürk, Y. (1996). *Uzaktan eğitim sistemlerinin tasarımı*. 1. Türkiye Uluslararası Uzaktan Eğitim Sempozyumunda sunulan bildiri, Ankara, Türkiye.

Başol, G. (2013). *Eğitimde ölçme ve değerlendirme* (Genişletilmiş 2. Baskı). Ankara: Pegem Akademi Yayıncılık.

Bayazıt, A. (2007). *Çevrimiçi sınavlar ve kâğıt-kalem sınavları arasındaki sınav süresi ve öğrenci başarım farklılıkları* (Yayımlanmamış yüksek lisans tezi). Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.

Baykul, Y. (2015). *Eğitimde ve psikolojide ölçme: klasik test teorisi ve uygulaması (3. Baskı)*. Ankara: Pegem Akademi Yayıncılık.

Bayyurt, Y. ve Akcan, S. (2014). *Türkiye'deki İngilizce öğretmeni yetiştirme programları üzerine düşünce ve öneriler*. 3. Ulusal Yabancı Dil Eğitimi Kurultayında sunulan bildiri, Boğaziçi Üniversitesi, İstanbul.

Berberoğlu, G. (2006). *Sınıf içi ölçme ve değerlendirme teknikleri*. İstanbul: Morpa Kültür Yayınları.

Brown, H. D. (2001). *Language assessment: Principles and classroom practices*. New York: Longman Pearson Education.

Brown, H. D. (2004), *Teaching by Principles: an interactive approach to language pedagogy*, Longman.

Brown, H. D. (1987), Principles of language learning and teaching (2nd Edition). Englewood Cliffs, New Jersey: Prentice Hall Regents.

Brown, J. D. (2005). *Testing in language programs*. McGraw-Hill College.

Brown, J. D., & Hudson, T. (2002). *Criterion-referenced language testing*. Cambridge: Cambridge University Press.

Büyüköztürk, Ş. (2016).Sosyal bilimler için veri analizi el kitabı (22. Basım). Ankara: Pegem Akademi Yayıncılık.

Campton, P. (2004). *A comparative analysis of online and paper-based assessment methods: A university case study*, paper presented at the Australian Education Conference, Newcastle, NSW.

Candrlic, S., Katić, M. A., & Dlab, M. H. (2014). *Online vs. paper-based testing: A comparison of test results*. paper presented in the 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), IEEE.

Carroll, J. & Sapon, S. (1959). *Modern language aptitude test* (MLAT). San Antonio, CA: Psychological Corporation.

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 83–118). Rowley, MA: Newbury House.

Cassady, J. C., & Gridley, B. E. (2005). The effects of online formative and summative assessment on test anxiety and performance. *The Journal of Technology, Learning and Assessment*, 4(1),4-30.

Chmiliar, l. (2010). Multiple-case designs. In A. J. Mills, G. Eurepas & E. Wiebe (Eds.), *Encyclopedia of case study research* (pp 582-583). USA: SAGE Publications.

Clark, J. L. D., & Lett J. (1988). A research agenda (Ed.by. Pardee lowe, Jr., Charles W. Stansfield) *Language in education: Theory and practice second language proficiency assessment: Current issues, language in education: Theory and Practice* (pp. 53-82). New Jersey: Prentice-Hall, Inc.

Cohen, R. J., & Swerdlik, M. E. (2013). *Psikolojik test ve değerlendirme: testlere ve ölçmeye giriş*. Nobel Akademik Yayıncılık.

Council of Europe (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, England: Cambridge University Press.

Council of Europe (2018). *Common European framework of reference for languages: Learning, teaching, assessment. Companion Volume with New Descriptors.* Cambridge, England: Cambridge University Press.

Coombe, C. (2018). *An A to Z of second language assessment: How language teachers understand assessment concepts*. London, UK: British Council.

Creswell, J. W. (2007). *Qualitative inquiry & research design: Choosing among five approaches* (2nd Ed.). USA: SAGE Publications.

Çakıcı, Y., Gemici, Ö. ve Özsevgeç, T. (2008). Fen ve teknoloji öğretiminde yeni yaklaşımlar. Ö. Taşkın (Ed.), *Fen ve teknoloji öğretiminde yapılandırmacı yaklaşım* içinde (s. 1-22). Ankara: Pegem Akademi Yayıncılık.

Çakır, İ. (2018). İngilizce öğretimi. A. Uzunöz ve V. Aktepe, (Ed.), *Özel öğretim yöntemleri* Cilt- 2 İçinde (1. Baskı), (s. 213-237). Ankara: Pegem Akademi.

Çelik, Z. (2006). *Web tabanlı otomasyon ve ölçme değerlendirme yönetim sistemi: KTU Fatih Eğitim Fakültesi örneği* (Yayımlanmamış yüksek lisans tezi). Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Trabzon.

Çepni, S. (2005). *Kuramdan uygulamaya fen ve teknoloji* (4.baskı). Ankara: Pegem Yayınları.

Çepni, S., Bayrakçeken, S., Yılmaz, A., Yücel, C., Semerci, Ç., Köse, E., Sezgin, F., Demircioğlu, G. ve Gündoğdu, K. (2011). Eğitimde ölçme ve değerlendirme. Karip, E. (Ed). *Ölçme ve Değerlendirme* içinde (2.baskı). Ankara: Pegem Akademi.

Çıkrıkçı-Demirtaşlı, R. N. (2014). *Öğrenme, öğretim ve değerlendirme arasındaki ilişkiler*. Ankara: Edge Akademi Yayın Dağıtım.

Çoruhlu, T. Ş., Nas, S. E. ve Çepni, S. (2009). Fen ve teknoloji öğretmenlerinin alternatif ölçme-değerlendirme tekniklerini kullanmada karşılaştıkları problemler: Trabzon örneği. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 6 (1), 122-141.

Çörekçioğlu, S. (2017). *A study on teacher perceptions and student attitudes towards the online English exams* (Unpublished master's thesis). Bahçeşehir University, Graduate School of Educational Sciences, İstanbul.

Damon, W. (2005). Fwd: Arresting Insights in Education. *lhomas Fordbarn Foundation*, 2(3), 1-5.

Davidson, F., & Lynch, B. K. (2008). *Testcraft: A teachers guide to writing and using language test specifications*. US: Yale University Press.

Davies, A. (1990). Operationalising uncertainty in language testing: An argument in favors of content validity, (Ed: John H. A.L. De Jong, -Douglas K. Stevenson) *Individualizing the assessment of language abilities, multilingual matters LTD* (pp.179-195). Clevedon Philadelphia,

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

Demirel, Ö. (1999). *Planlamadan değerlendirmeye öğretme sanatı*. Ankara: Pegem Akademi.

Doughty, C., Campbell, S., Bunting, M., Bowles, A., & Haarmann, H. (2007). *Technical report: The development of the High-Level Language Aptitude Battery*. College Park: University of Maryland Center for Advanced Study of Language.

Doughty, C., Campbell, S., Bunting, M., Mislevy, M., Bowles, A., & Koeth, J. (2010). *Predicting near-native L2 ability: The factor structure and reliability of Hi-LAB* (pp. 10–31). In M. T. Prior, Y. Watanabe, & S.-K. Lee (Eds.), Selected proceedings of the 2008 Second Language Research Forum: Exploring SLA perspectives, positions, and practices. Somerville, MA: Cascadilla Press.

East, M. (2008). *Dictionary use in foreign language writing exams: impact and implications*. PA: John Benjamins.

Eckes, T., & Grotjahn, T. (2006). A closer look at the construct validity of C-Tests. *Language Testing*, 23, 290–325.

Ekiz, D. (2009). *Bilimsel araştırma yöntemleri* (2. Baskı). Ankara: Anı Yayıncılık.

Emir, Ş. (2006). *E-öğrenmede sınav modelleri ve uygulaması* (Yayımlanmamış yüksek lisans tezi). İstanbul Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

Ercan, İ. ve Kan, İ. (2004). Ölçeklerde güvenirlik ve geçerlik. *Uludağ Üniversitesi Tıp Fakültesi Dergisi*, 30(3), 211-216.

Ercengiz, M. (2020). Dürtüsellik ile sosyal medya aşerme arasındaki ilişkide FOMO'nun aracılık etkisi. *Uluslararası Toplum Araştırmaları Dergisi*, *15*(22), 1291-1314.

Erdem, C. ve Eskimen, A. D. (2019). Türk dili ve edebiyatı öğretiminde dil bilinç düzeyinin gelişimi üzerine bir araştırma. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, *49*, 23-37.

Erkuş, A. (2006). *Sınıf öğretmenleri için ölçme ve değerlendirme kavramlar ve uygulamalar*. Ankara: Ekinoks Yayıncılık.

Fask, A., Englander, F., & Wang, Z. (2014). Do online exams facilitate cheating? An experiment designed to separate possible cheating from the effect of the online test taking environment. *Journal of Academic Ethics*, *12*(2), 101-112.

Field, A. (2009) Discovering statistics using SPSS (3rd Edition). London: Sage Publications Ltd.

Gaytan, J. (2005). Effective assessment techniques for online instruction. *Information Technology, Learning, and Performance Journal*, *23*(1), 25.

Gelen, İ. ve Beyazit, N. (2007). Eski ve yeni ilköğretim programları ile ilgili çeşitli görüşlerin karşılaştırılması. *Educational Administration: Theory and Practice*, *51*, 457-476.

Göçer, A. (2018). *Türkçe eğitiminde ölçme ve değerlendirme*. Ankara: Pegem Akademi.

Gronlund, N. E., & Waugh, C.K. (2009). *Assessment of student achievement*. New Jersey: Pearson Education Inc.

Gronlund, N. E. (1985) *Measurement and evaluation in teaching* (5th. Ed.). New York: Macmillan Co.

Güler, N. (2018). *Eğitimde ölçme ve değerlendirme* (12. baskı.). Ankara: Pegem Akademi.

Gümüş, B. (1977). *Eğitimde ölçme ve değerlendirme*. Ankara: Kalite Matbaası.

Gündoğmuş, İ. (2017*). Kağıt-kalem ve tablet ortamında gerçekleştirilen sınavlar için ölçme değişmezliğinin ve öğrenci görüşlerinin incelenmesi* (Yayımlanmamış yüksek lisans tezi). Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.

Güngör, A. (2001). *İlköğretim Okulları ikinci kademe din kültürü ve ahlak bilgisi dersi öğretmenlerinin ölçme ve değerlendirme sorunları* (Yayımlanmamış yüksek lisans tezi). Erciyes Üniversitesi, Sosyal Bilimler Enstitüsü, Kayseri.

Hancock, R.D., & Algozzine, B. (2006). *Doing case study research*. New York: Teachers College Press.

Harlen, W. (2005). Teachers' summative practises and assessment for learning tensions and synergies. *The Curriculum Journal*, *16*(2), 207-223.

Harmer, J. (2001).*The practice of English language teaching*. England: Pearson Education Limited.

Harmon, O. R., & Lambrinos, J. (2008). Are online exams an invitation to cheat? *The Journal of Economic Education*, *39* (2), 116-125.

Harrison, A. (1983). *A language testing handbook*. London: MacMillan Publishers Limited.

Hasselgreen, A. (2004). *Testing the spoken English of young Norwegians: A study of test validity and the role of 'smallwords' in contributing to pupils fluency, studies in language testing*. Cambridge: Cambridge University Press

Herzog, M. (1988). Issues in writing proficiency assessment, (Ed: Pardee lowe, Jr., Charles W. Stansfield) *Language in education: Theory and practice second language proficiency assessment: Current issues* (p. 149-175).New Jersey : Prentice-Hall, Inc.

Hidri, S. (2018). *Discrete point and integrative testing*. In the TESOL Encyclopedia of English Language Teaching (eds J. I. Liontas, T. International Association and M. DelliCarpini).

Howitt, D., & Cramer, D. (2011). Introduction to SPSS statistics in psychology: For version 19 and earlier. Pearson.

Hughes, A. (2003). *Testing for language teachers*. Cambridge: Cambridge University Press.

Hylton, K., Levy, Y., & Dringus, L. P. (2016). Utilizing webcam-based proctoring to deter misconduct in online exams. *Computers & Education*, *92*, 53-63.

İçten, T. (2006). *Uzaktan eğitim öğrencileri için web tabanlı çevrimiçi sınav sistemi uygulaması geliştirilmesi* (Yayımlanmamış yüksek lisans tezi). Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.

İşyar, Ö. Ö. (2017). *Sınıf öğretmenlerinin eğitimde drama kullanımına ilişkin yeterlilik algıları ve drama kavramına yönelik metaforları ile görüşlerinin incelenmesi* (Yayımlanmamış yüksek lisans tezi). Mersin Üniversitesi, Eğitim Bilimleri Enstitüsü, Mersin.

Jamil, M., Tariq, R. H., & Shami, P. A. (2012). Computer-based vs paper-based examinations: Perceptions of university teachers. T*urkish Online Journal of Educational Technology-TOJET*, *11*(4), 371-381.

Jeong, H. (2014). A comparative study of scores on computer-based tests and paper-based tests. *Behaviour & Information Technology*, *33*(4), 410-422.

JISC (2007). *Effective practice with e-assessment an overview of technologies, policies and practice in further and higher education*. 06.03.2021 tarihinde http://www.jisc.ac.uk/media/documents/themes/elearning/effpraceass ess.pdf adresinden erişilmiştir.

Jia, J., & He, Y. (2021). *The design, implementation and pilot application of an intelligent online proctoring system for online exams.* Interactive Technology and Smart Education.

Jung, I. Y., & Yeom, H. Y. (2009). Enhanced security for online exams using group cryptography. *IEEE transactions on Education*, *52*(3), 340-349.

Kan, A. (2017). Ölçmenin temel kavramları. H. Atılgan, (Ed.), *Eğitimde ölçme ve değerlendirme* (10. Baskı) içinde (s.19-42). Ankara: Anı Yayıncılık.

Kanatlı, F. (2008). Alternatif Ölçme Değerlendirme Teknikleri Konusunda Sınıf Öğretmenlerinin Görüşlerinin Değerlendirilmesi (Yayımlanmamış yüksek lisans tezi). Mustafa Kemal Üniversitesi, Sosyal Bilimler Enstitüsü, Hatay.

Karaağaç, G. (2005). *Dil, tarih ve insan*. Ankara: Akçağ Yayınları.

Karaca, E., Gömleksiz, M., Yurdabakan, İ., Çetin, B., Nartgün, Z., & Bıçak, B. (2014). *Eğitimde ölçme ve değerlendirme*. Nobel Yayınları.

Katz, L. G. 1993. *Dispositions as educational goals*. ERIC Digest. ERIC Document Reproduction Service: ED363454.

Kaya, Ü. ve Semerci, Ç. (2017). Yükseköğretimde bağıl ve mutlak değerlendirmeye ilişkin öğretim elemanlarının görüşleri. *Akademik Sosyal Araştırmalar Dergisi*, *5*(47),457-467.

Kesen, İ. (2019). 2023 *Eğitim vizyonu belgesinde ölçme ve değerlendirme*. Seta Perspektif.

Kılıç, G. B. (2006). *Yeni yaklaşımlar ışığında ilköğretim bilim öğretimi*. İstanbul: Morpa Yayıncılık.

Kılıç, S. (2015).Kappa test. *Journal of Mood Disorders*, *5*(3), 142-144.

Kırık, M. Y. (2008). *Yabancı dil olarak İngilizce öğretmenlerinin ölçme değerlendirme bağlamında tutum ve yaklaşımları* (Yayımlanmamış doktora tezi). İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.

Kip-Kayabaş, B. (2014). Dil öğretiminde online değerlendirme: Türkçe e-sertifika programı örneği. *JRET - Eğitim ve Öğretim Araştırmaları Dergisi*, 3 (3), 181-188.

Kuyumcu, A. ve Erdoğan, T. (2007). Öğrencilerimizi değerlendirme sürecine katılımı: Öz ve akran değerlendirme. *İlköğretim Eğitimci Dergisi*, *13*, 10-11.

Landis, J., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159-174.

Lett, J., Herzog, M., Keesling, W., Jackson, G., Funke, M., & Krol, M. (2004). Background to the defense language aptitude battery (DLAB) for issue paper authors. In D. M. Kenyon & D. MacGregor (Eds.), *Final report of the defense language aptitude battery II project* (pp. B1–B12). Monterey, CA: Defense Language Institute Foreign Language Center.

Linn, R. L., & Gronlund, N. E. (1995) *Measurement and assessment in teaching* (7th. Ed.). New Jersey: Prentice-Hall, Inc.

Lowe, J. P. (1988). The unassimilated history (Ed: Lowe. Jr., Pardee – Charles W. Stansfield) Second language proficiency assessment: Current issues, language in education: Theory and practice (pp.11-51), New Jersey: Prentive-Hall, Inc.

McKillup, S. (2012). Statistics explained: An introductory guide for life scientists (Second edition). United States: Cambridge University Press.

Mcnamara, T. (2000). Language testing. *International Journal of Applied Linguistics*, 16 (3), 424-434.

MEB. (2005). *İlköğretim okulu ders programları ve öğretim kılavuzu (1-5. Sınıflar)*. Ankara: MEB Yayınları.

MEB. (2006). *EARGED ilköğretim 6. sınıf pilot uygulama sonuçlarının değerlendirilmesi*. Ankara: MEB Yayınları.

MEB. (2017a). *Ortaöğretim İngilizce dersi (9, 10, 11, 12, Sınıflar) öğretim programı*. http://mufredat.meb.gov.tr/ProgramDetay.aspx?PID=342file:///F:/201812020472656-OGM%20INGILIZCE%20PRG%2020.012018.pdf (Erişim Tarihi: 15.03. 2021).

MEB. (2017). *Öğretmenlik mesleği genel yeterlikleri*. Öğretmen Yetiştirme ve Geliştirme Genel Müdürlüğü, Ankara.

Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass Publishers.

Morgan, C., & O'Reilly, M. (1999). *Assessing open and distance learners*. London: KoganPage.

Nartgün, Z. (2014). Duyuşsal nitelikler ve ölçülmesi. S. Erkan & M. Gömleksiz (Ed.), *Eğitimde ölçme ve değerlendirme* içinde (ss. 144-196). Ankara: Nobel Akademik Yayıncılık.

Nartgün, Z. (2007). Aynı puanlar üzerinden yapılan mutlak ve bağıl değerlendirme uygulamalarının notlarda farklılık oluşturup oluşturmadığına ilişkin bir inceleme. *Ege Eğitim Dergisi*, 8(1), 19-40.

Navruz, M. (2011). *Bilgisayar ortamında yapılan sınavların akademik başarıya etkisinin incelenmesi* (Yayımlanmamış yüksek lisans tezi). Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya.

Nitko, A. ve Brookhart, S. M. (2016). *Öğrencilerin eğitsel değerlendirmesi*. (Bıçak, B., Bahar, M. & Özel, S., Çev.) Ankara: Nobel Akademik Yayıncılık.

Odacı, M.M. (2019). *Bilgisayar destekli ölçme platformunun bir ölçme aracı olarak kullanımına yönelik İngilizce öğretmen adaylarının tutumlarının ortaya çıkarılması ve bilişsel yüklerinin farklı değişkenler açısından incelenmesi* (Yayımlanmamış yüksek lisans tezi). Mersin Üniversitesi, Eğitim Bilimleri Enstitüsü, Mersin.

Oller, J.W., & London, Jr. (1979). *Language tests at school: A pragmatic approach*. London, England: Longman.

Ornstein, A., & Hunkins, F. (2009) Curriculum design. *In curriculum: Foundations, principles and issues* (5th Ed.), pp. 181-206. Boston, MA: Pearson/Allyn and Bacon.

Özbek, Ö.Y. (2017). Ölçme araçlarında bulunması istenen nitelikler, S. Tekindal (Ed.), *Eğitimde ölçme ve değerlendirme* içinde (s. 41-86). Ankara: Pegem Akademi Yayıncılık, 2017.

Özçelik, D. A. (2010). *Ölçme ve değerlendirme*. Ankara: Pegem Akademi.

Özel, S. (2006). *İnternet tabanlı ölçme uygulamalarının başarı puanlarını yordama gücüne ilişkin bir çalışma* (Yayımlanmamış yüksek lisans tezi). Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.

Özmen, K.S. ve Balçıkanlı, C. (2012). Yabancı dil eğitiminde web tabanlı ölçme ve değerlendirme A. Sarıçoban ve Z.M. Tavil (Eds.), *Yabancı dil Öğretiminde Öğretim teknolojileri ve materyal tasarımı* içinde (ss229-251). Ankara: Anı Yayıncılık.

Özturan, T. (2016). *The impact of computer assisted assessment on the exam success and attitudes of prospective English teachers* (Unpublished master's thesis). Hacettepe University, Graduate School of Educational Sciences, Ankara.

Özyürek, A. (2016). Dramatik etkinlikler ve drama ile ilgili temel kavramlar. E. Ömeroğlu (Ed.), *Okul öncesi eğitimde dramatik etkinliklerden dramaya teoriden uygulamaya* içinde (7. baskı, ss. 1-6) Ankara: Eğiten.

Palloff, R. M., & Pratt, K. (2003). *The virtual student: A profile and guide*. San Francisco: Jossey - Bass.

Palloff, R. M.,& Pratt, K. (2009). *Assessing the online learner*. San Francisco: Jossey - Bass.

Pamukçu, A. (2018). *A case study in an English preparatory school in Turkey: do online practice tests contribute to success in in-class exams?* (Unpublished master's thesis). Bahçeşehir University, Graduate School of Educational Sciences, İstanbul.

Perkins, D. N., & Tishman, S. (1998).*Dispositional aspects of intelligence*. Paper presented at the 1998 Winter Symposium of the College of Education, Arizona State University, Tempe.

Peterson, C. R., & Al-Haik, A. R. (1976). The development of the defense language aptitude battery (DLAB). *Educational and Psychological Measurement*, *36*, 369–80.

Polat, T. (2006). Okur odaklı bir yaklaşımla yazın eğitimi. *Millî Eğitim Dergisi*, *34* (169), 1-7.

Raths, J. (2001). Teachers' beliefs and teaching beliefs. *Early Childhood Research ve Practice*, *3*(1),354-395.

Rogier, D. (2014). Assessment literacy: Building a base for better teaching and learning. English Teaching Forum, 52(3), 2–13.

Semerci, Ç. (2015). Eğitimde ölçme ve değerlendirme. E. Karip (Ed.), *Ölçme ve değerlendirme* (7. Baskı) içinde (s.2-15). Ankara: Pegem Akademi.

Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. Pearson Education Limited.

Shuey, S. (2002). Assessing online learning in higher education. *Journal of Instruction Delivery Systems*, *16* (2), 13-18.

Sönmez, V. (2015). Eğitim felsefesi. Ankara: Anı Yayıncılık.

Stake, R. R. (2005). Case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (3rd Ed.). London: Sage.

Still, M. L., & Still, J. D. (2015). Contrasting traditional in-class exams with frequent online testing. *Journal of Teaching and Learning with Technology*, *4*(2), 30-40.

Stowell, J. R., & Bennett, D. (2010). Effects of online testing on student exam performance and test anxiety. *Journal of Educational Computing Research*, *42*(2), 161-171.

Şahin, Y. (2018). *Yabancı Dilde Ölçme ve Değerlendirme*. Ankara: Pegem Akademi.

Şahin, Y.(2013). Dil, Ed. Y. Şahin, *Farklı boyutlarıyla yabancı dil öğrenimi ve öğretimi* içinde(s. 10-16).Konya: Eğitim Kitabevi.

Şimşek, İ. (2013). *Eğitimde ölçme ve değerlendirme kriterlerine uygun web tabanlı uzman sınav sistemi geliştirilmesi* (Yayımlanmamış doktora tezi). İstanbul Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

Tafazoli, D. (2018). *An a to z of second language assessment: How language teachers understand assessment*. Dictionary Contribution.

Tan, Ş. (2006). *Öğretimi planlama ve değerlendirme*. Ankara: Pegem Akademi.

Tan, Ş. (2014). *Öğretimde ölçme ve değerlendirme KPSS el kitabı*. Ankara: Pegem Akademi.

Tavşancıl, E. ve Aslan, E. (2001). *Sözel, yazılı ve diğer materyaller için içerik analizi ve uygulama örnekleri*. İstanbul: Epsilon Yayınevi.

Tekin, H. (2000). *Eğitimde ölçme ve değerlendirme*. Ankara: Yargı Yayınları.

Tekindal, S. (2016). *Okullarda ölçme ve değerlendirme yöntemleri* (5. Baskı). Ankara: Nobel Akademik Yayıncılık.

Turgut, M. F. (1997). *Eğitimde ölçme ve değerlendirme metotları*. Ankara: Nuve Matbaası.

Turgut, M. F., & Baykul, Y. (2015). *Eğitimde ölçme ve değerlendirme*. Ankara: Pegem Akademi.

Turner, C. E. (2001).The need for impact studies of L2 performance testing and rating: Identifying areas of potential consequences at all levels of the testing cycle, (Ed: C. Elder –A. Brown –E. Grove –K. Hill –N. Iwashita –T. Lumley –T. McNamara –K. O'Loughlin) *Experimenting with uncertainty, studies in language testing* (pp.138-149) Cambridge University Press.

Türnüklü, A. (2000). Eğitimbilim araştırmalarında etkin olarak kullanılabilecek nitel bir araştırma tekniği: Görüşme. *Kuram ve Uygulamada Eğitim Yönetimi*, *24*, 543-559.

Urgun, A. (2019). *Ortaöğretim öğrencilerine yönelik mobil destekli ve kâğıt kalem testlerinin karşılaşırılması* (Yayımlanmamış yüksek lisans tezi). Akdeniz Üniversitesi, Eğitim Bilimleri Enstitüsü, Antalya.

Uysal, M. Öztürk, H. ve Döş, İ. (2015). *Eğitimde ölçme ve değerlendirme* (2. baskı). Ankara: Nobel Akademik Yayıncılık.

Vallette, R. M. (1977). *Modern language testing* (2nd Ed.) Harcourt Brace Jovanovich, Inc.

Vazquez, J. J., Chiang, E. P., & Sarmiento-Barbieri, I. (2021). Can we stay one step ahead of cheaters? A field experiment in proctoring online open book exams. *Journal of Behavioral and Experimental Economics,* 90, 101653.

Wall, D. (2005). *The impact of high-stakes examinations on classroom teaching : A case study using insights from testing and innovation theory*. Cambridge: Cambridge University Press.

Wall, D. –Alderson, J. C. (1996). Examining washback: The Sri Lankan impact study. (Ed: Alister Cumming –Richard Berwick) *Validation in language testing* (pp. 194-221). Multilingual Matters.

Wallace, P., & Clariana, R. B. (2005). Gender differences in computer-administered versus paper-based tests. *International Journal of Instructional Media*, *32*(2), 171.

Yağcı, M. (2012). *Yeni bir sınav modelinin geliştirilmesi ve kâğıt kalem testleri ile karşılaştırılması* (Yayımlanmamış doktora tezi). Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Sakarya.

Yaşar, M. (2017). Ölçme ve değerlendirme ile ilgili temel kavramlar. S. Tekindal, (Ed.), *Eğitimde ölçme ve değerlendirme* (5. Baskı) içinde (s.9-40). Ankara: Pegem Akademi.

Yayla, R. G. (2012). *Fen ve teknoloji öğretmenlerinin alternatif ölçme değerlendirme yöntem teknikleri ve uygulamaları hakkındaki görüş ve düşüncelerinin belirlenmesi* (Yayımlanmamış yüksek lisans tezi). Karadeniz Teknik Üniversitesi, Eğitim Bilimleri Enstitüsü, Trabzon.

Yıldırım, A., & Şimşek, H. (2018). *Sosyal bilimlerde nitel araştırma yöntemleri*. Ankara: Seçkin.

Yılmaz İnce, E. (2016). *Web ortamındaki yazılı sınavların doğal dil işleme yöntemleri ile değerlendirilmesi* (Yayımlanmamış doktora tezi). Süleyman Demirel Üniversitesi, Fen Bilimleri Enstitüsü, Isparta.

Yılmaz, İ. (2018). Yabancılara Türkçe deyimlerin öğretiminde tamamlayıcı ölçme ve değerlendirme yöntemlerinin kullanılması. *Kastamonu Eğitim Dergisi*, *26* (5), 1623-1641.

Yılmaz, K. (2011). Sosyal bilgiler öğretiminde ölçme ve değerlendirme. C. Öztürk. (Ed.). *Sosyal bilgiler öğretimi demokratik vatandaşlık eğitimi* (2. baskı) içinde (ss.431-515). Ankara: Pegem Akademi.

Yiğit, M. (2017). *Kültürel etkileşim bağlamında yabancı dil olarak Türkçe öğretimi "Arnavutlukta kültürel etkileşim desenleriye"*. İstanbul: Hiper Yayın.

Yin, R. K. (2009). *Case study research: Design and methods* (4th Ed.). Thousand Oaks, CA: Sage.

Yin, R. K. (1994). Discovering the future of the case study. Method in evaluation research. Evaluation practice, 15(3), 283-290.

Appendix-1. *Normality Results of Quantitative Data*

Dimension	Variable	K-S	Skewness	Kurtosis	Skewness (Z)	Kurtosis (Z)
Success in face-to-face exams	Female	0,000	-0,13	-0,22	0,07	-0,28
	Male	0,000	0,03	-1,05	-0,27	-1,39
Success in online exams	Female	0,000	-0,59	0,45	1,49	0,52
	Male	0,000	-0,18	-0,76	0,18	-0,41
Stress in face-to-face exams	Female	0,000	-0,65	-0,29	0,80	-0,77
	Male	0,000	-0,03	-1,58	1,24	0,24
Stress in online exams	Female	0,000	-0,39	-0,83	0,76	1,95
	Male	0,000	0,22	-1,04	-0,42	-0,16
Success in face-to-face exams	18	0,000	-0,75	0,72	0,07	-0,28
	19	0,000	0,42	-0,62	-0,27	-1,39
	20	0,000	0,14	-1,18	1,49	0,52
	21	0,035	0,46	0,52	0,18	-0,41
Success in online exams	18	0,000	-0,70	1,37	0,80	-0,77
	19	0,000	-0,17	0,09	1,24	0,24
	20	0,004	-0,40	-0,83	0,76	1,95
	21	0,002	-1,16	2,47	-0,42	-0,16
Stress in face-to-face exams	18	0,000	-0,67	-0,92	0,07	-0,28
	19	0,000	-0,27	-1,26	-0,27	-1,39
	20	0,000	-0,59	-1,22	1,49	0,52
	21	0,008	-1,16	1,54	0,18	-0,41
Stress in online exams	18	0,004	-0,26	-0,58	0,80	-0,77
	19	0,000	-0,32	-0,67	1,24	0,24
	20	0,000	0,19	-1,73	0,76	1,95
	21	0,042	-0,22	-1,57	-0,42	-0,16
Success in face-to-face exams	Elementary	0,001	0,37	-2,80	0,07	-0,28
	Pre-intermediate	0,000	-0,04	-1,18	-0,27	-1,39
	Intermediate	0,000	0,19	-0,29	1,49	0,52
	Upper-intermediate	0,002	-0,37	0,41	0,18	-0,41
	Advanced	0,000	0,03	-0,24	0,80	-0,77
Success in online exams	Elementary	0,016	-0,65	-1,70	1,24	0,24
	Pre-intermediate	0,000	-0,26	0,30	0,76	1,95
	Intermediate	0,000	-0,24	-0,12	-0,42	-0,16
	Upper-intermediate	0,000	0,03	-0,24	0,07	-0,28
	Advanced	0,006	-0,42	-0,58	-0,27	-1,39
Stress in face-to-face exams	Elementary	0,020	-1,12	0,27	1,49	0,52
	Pre-intermediate	0,000	-0,31	-1,46	0,18	-0,41
	Intermediate	0,000	-0,85	-0,25	0,80	-0,77
	Upper-intermediate	0,023	-0,59	0,33	1,24	0,24
	Advanced	0,018	0,03	-1,32	0,76	1,95
Stress in online exams	Elementary	0,020	1,12	0,27	0,07	-0,28
	Pre-intermediate	0,001	-0,51	-0,59	-0,27	-1,39
	Intermediate	0,023	-0,06	-1,30	1,49	0,52
	Upper-intermediate	0,057	0,17	-0,86	0,18	-0,41
	Advanced	0,018	-0,35	-0,96	0,80	-0,77
Success in face-to-face exams	Novice	0,003	-0,09	-0,90	0,07	-0,28
	Medium	0,000	-0,14	-0,07	-0,27	-1,39
	Advanced	0,000	0,10	-1,10	1,49	0,52
	Expert	0,086	-0,60	-0,35	0,18	-0,41
Success in online exams	Novice	0,154	-0,12	-0,15	0,80	-0,77
	Medium	0,000	0,19	-0,38	1,24	0,24
	Advanced	0,000	0,17	-0,08	0,76	1,95
	Expert	0,002	-2,12	4,74	-0,42	-0,16
Stress in face-to-face exams	Novice	0,000	-0,71	-1,26	0,07	-0,28
	Medium	0,000	-0,42	-0,82	-0,27	-1,39
	Advanced	0,000	-0,64	-0,83	1,49	0,52
	Expert	0,013	-1,57	1,97	0,18	-0,41
Stress in online exams	Novice	0,000	-1,41	1,79	0,80	-0,77
	Medium	0,000	-0,11	-0,89	1,24	0,24
	Advanced	0,000	0,71	-0,38	0,76	1,95
	Expert	0,224	0,31	-1,83	-0,42	-0,16

Öğretmenlerin ve öğrencilerin gözünden yabancı dil öğretiminde çevrim içi (online) sınavların değerlendirilmesi

Değerli katılımcı,

Bu çalışmada, öğretmenlerin ve öğrencilerin gözlerinden çevrim içi sınavların değerlendirilmesini amaçlamaktadır.

Araştırma kapsamında toplanacak veriler, sadece bilimsel amaçla kullanılacak olup hiçbir kişisel veriniz herhangi bir nedenle paylaşılmayacaktır.

Soruları içtenlikle cevaplamanız araştırmanın bulgularına önemli katkı sağlayacaktır.

Araştırmaya katkılarınızdan dolayı şimdiden teşekkür ederim.

Öğr. Gör. Kerime Asma

kerimeasmaa@gmail.com

Demografik Bilgiler

Bu bölümde, demografik bilgilerinizi tespit etmeye yönelik üç soru yer almaktadır.

Cinsiyetiniz:

Yaşınız:

Eğitim gördüğünüz (yeni tamamladığınız) İngilizce seviyesi (elementary, pre-intermediate, intermediate, upper intermediate vb.)

Elementary	Pre-intermediate	Intermediate	Upper-intermediate	Advanced

Çoktan Seçmeli Sorular

Bu bölümde, altı çoktan seçmeli soru yer almaktadır.

Yüz yüze sınavlarda başarı durumunuz:

Çok kötü	1	2	3	4	5	Çok iyi

Çevrim içi sınavlarda başarı durumunuz:

Çok kötü	1	2	3	4	5	Çok iyi

Yüz yüze sınavlarda stres seviyeniz:

Hiç	1	2	3	4	5	Aşırı

Çevrim içi sınavlarda stres seviyeniz:

Hiç	1	2	3	4	5	Aşırı

Bilgisayar kullanım beceriniz:

Acemi	Orta	İleri	Uzman

İnternet kullanım geçmişiniz:

1 yıldan az	1-3 yıl	3-5 yıl	5 yıldan fazla

Açık Uçlu Sorular

Bu bölümde, dokuz açık uçlu soru yer almaktadır.

1. Çevrim içi sınavların ne tür avantajları / dezavantajları olduğunu düşünüyorsunuz? Lütfen açıklayınız.
2. Çevrim içi sınavlarda karşılaştığınız genel problemler nelerdir? Lütfen belirtiniz.
3. Sizce çevrim içi sınavlarda değiştirilmesi / geliştirilmesi gereken bölüm / bölümler var mıdır? Varsa bunlar nelerdir? Lütfen nedenini açıklayınız.
4. Çevrim içi sınavlarda yüz yüze sınavlara oranla ekstra süre verilmesi gerektiğini düşünüyor musunuz? Neden?
5. Çevrim içi sınavlarda gerçek performansınızı gösterdiğinizi düşünüyor musunuz? Lütfen açıklayınız.
6. Çevrim içi sınavlarda en çok zorlandığınız beceri hangisidir? Neden?
7. Çevrim içi sınavlarda yaşadığınız problemler sınavın çevrim içi olmasından mı kaynaklanıyor, yoksa genel olarak yaşadığınız problemler mi? Lütfen açıklayınız.
8. Bir tercih hakkınız olsaydı, çevrim içi sınavları mı yoksa yüz yüze sınavları mı tercih ederdiniz? Neden?
9. Dil becerisinin/becerilerinin (okuma - dinleme - konuşma - yazma) sınavı çevrim içi olarak devam ettirilecek olsa, bu becerinin/becerilerin hangisi olmasını isterdiniz? Lütfen açıklayınız.

Appendix-3. *Online Form for Instructors*

Examining online testing practices in foreign language education through the lens of instructors and students

Dear participant,

This study aims at examining online testing practices in foreign language education through the lens of teachers and students.

The data collected within the scope of the research will be kept confidential and used solely for scientific purposes.

Your sincere answer to the questions will significantly contribute to the findings of the research.

Thank you in advance for your contribution to the research.

Inst. Kerime Asma

kerimeasmaa@gmail.com

Demographic Information

Gender:

Age:

Educational Background:

Teaching experience in this institution (year):

Teaching experience (Overall):

Computer skills:

Basic	Medium	Advanced	Expert

Active internet usage background:

Less than one year	1-3 years	3-5 years	More than 5 years

Distance/online education experience status (prior to COVID-19 pandemic):

YES	NO

If your answer is "YES" for the previous question, for how long? :

Open-ended questions

1. How do you find online testing practices in your institution? Do you think the existing practices are practical if the distance education continues? Please explain.

2. Are there any deficiency / deficiencies that you have detected or experienced in online testing practices? If any, what are your suggestions for eliminating them?

3. Are there any particular skill/skills you have difficulty measuring in online testing practices? If any, what kind of problems do you have? Please explain.

4. What are the common problems you experience with online testing practices? Please explain.

5. In your opinion, what are the advantages and disadvantages of online testing? Please explain.

6. If you were to develop an alternative model to existing online testing practices, what kind of model would you design? Why? Please explain.

ÖZ GEÇMİŞ

Kişisel Bilgiler

Adı ve Soyadı   : Kerime ASMA

Doğum Yeri ve Tarihi   : Antalya / 19.03.1990

Eğitim Durumu

Lisans Öğrenimi   : Akdeniz Üniversitesi, Eğitim Fakültesi

İngilizce Öğretmenliği (2010-2015)

Bildiği Yabancı Diller   : İngilizce, İspanyolca

İş Deneyimi

Çalıştığı Kurumlar   : Antalya Bilim Üniversitesi (2017-halen çalışıyor)

İletişim

E-posta   : kerimeasmaa@gmail.com

Tarih   :16.08.2021

BİLDİRİM

Hazırladığım tezin tamamen kendi çalışmam olduğunu ve her alıntıya kaynak gösterdiğimi taahhüt eder, tezimin kâğıt ve elektronik kopyalarının Akdeniz Üniversitesi Eğitim Bilimleri Enstitüsü arşivlerinde aşağıda belirttiğim koşullarda saklanmasına izin verdiğimi onaylarım.

☐ Tezimin tamamı her yerden erişime açılabilir.

☐ Tez sadece Akdeniz Üniversitesi yerleşkelerinden erişime açılabilir.

☑ Tezimin/Raporumun 1 (bir) yıl süreyle erişime açılmasını istemiyorum. Bu sürenin sonunda uzatma için başvuruda bulunmadığım takdirde, tezimin tamamı her yerden erişime açılabilir.

16.08.2021

Kerime ASMA

İNTİHAL RAPORU

EXAMINING ONLINE TESTING PRACTICES IN FOREIGN
LANGUAGE EDUCATION THROUGH THE PERCEPTIONS OF
INSTRUCTORS AND STUDENTS: A CASE STUDY