



HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Department of Foreign Languages Education

English Language Education Program

UNIVERSITY ENGLISH PREPARATORY SCHOOL STUDENTS' ATTITUDES TOWARDS
ASSESSMENT AND EVALUATION INSTRUMENTS

Ayça İrem YILMAZ

Master's Thesis

Ankara, 2022



With leadership, research, innovation, high quality education and change,

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



HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Department of Foreign Languages Education

English Language Education Program

UNIVERSITY ENGLISH PREPARATORY SCHOOL STUDENTS' ATTITUDES TOWARDS
ASSESSMENT AND EVALUATION INSTRUMENTS

ÜNİVERSİTE İNGİLİZCE HAZIRLIK SINIFI ÖĞRENCİLERİNİN ÖLÇME ve
DEĞERLENDİRME ARAÇLARINA YÖNELİK TUTUMLARI

Ayça İrem YILMAZ

Master's Thesis

Ankara, 2022

Acceptance and Approval

To the Graduate School of Educational Sciences,

This thesis, prepared by **AYÇA İREM YILMAZ** and entitled “University English Preparatory School Students’ Attitudes Towards Assessment and Evaluation Instruments” has been approved as a thesis for the Degree of **Master** in the **Program of English Language Education** in the **Department of Foreign Languages Education** by the members of the Examining Committee.

Chair Prof. Dr. İsmail Hakkı MİRİCİ

Member (Supervisor) Asst.Prof.Dr. İsmail Fırat ALTAY

Member Asst. Prof. Dr. Deren Başak
Akman YEŞİLEL

This is to certify that this thesis has been approved by the aforementioned examining committee members on 17/06/2022 in accordance with the relevant articles of the Rules and Regulations of Hacettepe University Graduate School of Educational Sciences, and was accepted as a **Master’s Thesis** in the **Program of English Language Education** by the Board of Directors of the Graduate School of Educational Sciences from/...../.....

Prof. Dr. Selahattin GELBAL

Director of Graduate School of Educational Sciences

Abstract

This study aims to investigate the attitudes of English language learners studying at a university preparatory school toward assessment and evaluation instruments. The study's population was A1-level students enrolled in a preparatory class during the fall semester of the 2021-2022 academic year. Both quantitative and qualitative data were collected for this study. The quantitative data was gathered through the questionnaire prepared by the researchers. For the qualitative part, a semi-structured interview form with 4 main questions and 3 sub-questions was employed.

The quantitative findings revealed that rather than having a mostly one-sided attitude toward testing instruments, students had moderate positive and negative attitudes. In the anxiety items of the questionnaire, there is a significant relationship between gender and student attitudes; however, there is no such relationship in the enjoyment items. In both sections of the questionnaire, the results revealed that an academic major makes a significant difference in students' attitudes.

The qualitative findings demonstrated that linguistic challenges and paired tasks are the causes of negative attitudes, and a sense of accomplishment is the cause of positive attitudes for the speaking assessment. Positive attitudes toward quizzes are related to the lower weighting of quizzes and the shorter test scope, negative attitudes toward mid-terms are due to the weighting of them and the bigger scope. Participants have a highly positive attitude toward the writing portfolio, and the most visible reason is receiving feedback from the teacher. Finally, participants have a slightly positive and noticeably negative attitude toward the online components, implying that they generate discontent and boredom.

Keywords: testing, assessment and evaluation instruments, attitude, anxiety, enjoyment.

Öz

Bu araştırma, üniversite İngilizce hazırlık sınıfı öğrencilerinin ölçme ve değerlendirme araçlarına yönelik tutumlarını araştırmayı amaçlamaktadır. Araştırmanın evreni, 2021-2022 akademik yılının güz döneminde eğitim almakta olan İngilizce başlangıç seviyesindeki (A1) öğrenciler oluşturmaktadır. Bu çalışma için hem nicel hem de nitel veriler toplanmıştır. Nicel veriler, araştırmacılar tarafından hazırlanan anket aracılığıyla toplanmıştır. Nitel kısımda 4 ana soru ve 3 alt sorudan oluşan yarı yapılandırılmış görüşme formu kullanılmıştır.

Nicel bulgular, ölçme ve değerlendirme araçlarına yönelik çoğunlukla tek taraflı bir tutuma sahip olmaktan ziyade, öğrencilerin orta derecede hem olumlu hem de olumsuz tutumlara sahip olduklarını ortaya koymuştur. Anketin kaygı maddelerinde öğrenci tutumları cinsiyete göre anlamlı değişiklik gösterirken, keyif alma maddelerinde böyle bir durum yoktur. Anketin her iki bölümünde de sonuçlar, öğrencilerin okuyacakları bölümlerin, tutumlarında anlamlı farklılık yarattığını ortaya koymaktadır.

Nitel bulgular, dilsel zorlukların ve partnerli çalışma gerektiren konuşma sınavlarının olumsuz tutumların nedenleri olduğunu ve konuşma sınavlarına ilişkin olumlu tutumların nedeninin ise başarı duygusu olduğunu göstermiştir. Kısa sınavlara (quizzes) yönelik olumlu tutumlar, bu sınavın ağırlığının öğrencilerin ortalaması üzerinde etkisinin az olması ve test kapsamının daha kısa olmasıyla ilgiliyken, vizelere yönelik olumsuz tutumlar ise bunların ortalamaya etkisinin fazla olması ve daha büyük kapsamlı olması ile ilgilidir. Katılımcıların yazma dosyasına (the writing portfolio) karşı oldukça olumlu bir tutumu vardır ve en görünür neden öğretmenden geri bildirim almaktır. Son olarak, katılımcıların, hoşnutsuzluk yarattıklarını ima ettikleri çevrimiçi bileşende yer alan alıştırmalar hakkında biraz olumlu, fakat belirgin şekilde olumsuz bir algıya sahip oldukları görülmüştür.

Anahtar sözcükler: ölçme, ölçme ve değerlendirme araçları, tutum, endişe, keyif alma.

Acknowledgements

First of all, I would like to express my deepest gratitude to my thesis advisor Asst. Prof. Dr. İsmail Fırat ALTAY for his supervision, guidance, encouragement, and immense knowledge throughout this research study, which has kept me motivated.

My sincere thanks also go to the committee of my thesis, Prof. Dr. İsmail Hakkı MİRİCİ, and Asst. Prof. Dr. Deren Başak Akman YEŞİLEL for providing constructive feedback on my thesis.

I'd also like to thank the university where I conducted the study, as well as my colleagues who helped me collect data. I'd also like to thank all of the students who took the time to complete the questionnaire and participate in the interviews.

My special thanks also go to my friend Kenan ACAROL for his constructive feedback and support for the data analysis. I owe much to him for his friendly comments, guidance, and support.

Last but not least, I would like to acknowledge my deep gratitude to my family who always supports me, and all people who supported me during this challenging process with their advice and support.

Finally, I owe the greatest thanks and the deepest gratitude to my beloved husband, M. Utku YILMAZ for his endless understanding, love, and belief in me.

Table of Contents

Acceptance and Approval	ii
Abstract	iii
Öz	iv
Acknowledgements	v
List of Tables	viii
List of Figures	ix
Symbols and Abbreviations	x
Chapter 1 Introduction	1
Statement of the Problem	1
Aim and Significance of the Study	5
Research Questions	6
Assumptions	7
Limitations	8
Definitions	9
Chapter 2 Theoretical Basis of Research and Literature Review	10
Language Assessment and Testing in General	10
The Attitude Concept	28
Anxiety in Language Learning	31
Positive Psychology and Foreign Language Enjoyment	39
Recent Studies	42
Chapter 3 Methodology	48
Types of Research	48
Setting and Participants	49
Data Collection	53
Instruments	53
Data Analysis	59

Chapter 4 Findings, Comments and Discussion	62
Quantitative Data Analysis	62
Qualitative Data Analysis	74
Chapter 5 Conclusion, Pedagogical Implications and Suggestions	101
Conclusion	101
Pedagogical Implications	104
Suggestions	108
References	110
APPENDIX-A: METU Human Subjects Ethics Committee Institutional Approval Form	cxxii
APPENDIX-B: The Questionnaire	cxxiii
APPENDIX-C: Semi-Structured Interview Form.....	cxxv
APPENDIX-D: Ethics Committee Approval	cxxvi
APPENDIX-E: Declaration of Ethical Conduct	cxxvii
APPENDIX-F: Thesis/Dissertation Originality Report	cxxviii
APPENDIX-G: Yayımlama ve Fikrî Mülkiyet Hakları Beyanı.....	cxxix

List of Tables

Table 1 <i>Descriptive Statistics of the Participant Students in the Quantitative Part</i>	51
Table 2 <i>Descriptive Statistics of the Participant Students in the Qualitative Part</i>	52
Table 3 <i>Questionnaire Items</i>	55
Table 4 <i>Reliability Analysis of the Questionnaire</i>	56
Table 5 <i>Data Collection Instruments</i>	58
Table 6 <i>Test of Normality Results of the Questionnaire</i>	59
Table 7 <i>Required Sample Size for the Research</i>	60
Table 8 <i>Analysis of the Average Mean Scores for the Anxiety and Enjoyment Items in the Questionnaire</i>	63
Table 9 <i>Descriptive Statistics of the Anxiety and Enjoyment Items in the Questionnaire</i>	64
Table 10 <i>Descriptive Statistics of Each Instrument in the Anxiety Section of the Questionnaire</i>	67
Table 11 <i>Descriptive Statistics of Each Instrument in the Enjoyment Section of the Questionnaire</i>	69
Table 12 <i>Independent T-test Results for Gender</i>	70
Table 13 <i>Independent T-test Results for Academic Major</i>	72

List of Figures

Figure 1 <i>Principles of Language Assessment</i>	23
Figure 2 <i>The Nature of Anxiety</i>	32
Figure 3 <i>Components of Foreign Language Anxiety</i>	34
Figure 4 <i>The Research Design of the Study (Adapted From Creswell, 2009)</i>	49
Figure 5 <i>Attitudes Towards Speaking Assessment</i>	75
Figure 6 <i>Linguistic Difficulties</i>	77
Figure 7 <i>Task Type- Paired Tasks</i>	78
Figure 8 <i>Fear of Negative Evaluation</i>	79
Figure 9 <i>Suggestions for Reducing the Effects of Negative Attitudes Toward Speaking Assessment</i>	80
Figure 10 <i>Reasons for Positive Attitudes Toward Speaking Assessment</i>	81
Figure 11 <i>Attitudes Towards Quizzes and Mid-term Exams</i>	84
Figure 12 <i>Reasons for Positive Attitudes Toward Quizzes</i>	86
Figure 13 <i>The Reasons for Negative Attitudes Toward Mid-term Exams and Suggestions</i>	87
Figure 14 <i>Attitudes Toward the Writing Portfolio</i>	90
Figure 15 <i>Reasons for Positive Attitudes Toward the Writing Portfolio</i>	90
Figure 16 <i>The Reasons for Negative Attitudes Toward the Writing Portfolio and Suggestions</i>	92
Figure 17 <i>Attitudes Toward the Online Components</i>	95
Figure 18 <i>Reasons for Positive Attitudes Toward the Online Components</i>	96
Figure 19 <i>The Reasons of Negative Attitudes Toward the Online Components</i> ..	98

Symbols and Abbreviations

CEFR: Common European Framework of Reference for Languages

EFL: English as Foreign Language

FLA: Foreign Language Anxiety

FLE: Foreign Language Enjoyment

L2: Second/Foreign language



Chapter 1

Introduction

This chapter establishes background about the significance of assessment and evaluation of English as a foreign language learning (EFL). The purpose of assessment and evaluation, the importance of them in the teaching and learning process, the common types of assessment and evaluation tools, and their roles in English language teaching are addressed in the statement of the problem. Afterward, the aims and significance of the study are presented. Subsequently, research questions, assumptions and limitations are displayed. Finally, this chapter defines the definitions that are used throughout the study.

Statement of the Problem

As teaching is the focus of the teaching-learning process, testing is also an indispensable part of this process. The utmost goal of these core concepts is to improve student learning. Since teaching and testing are strongly related to each other and are built upon each other's data, both of them provide necessary information to all the stakeholders to guide the instruction process seamlessly. Teachers' primary goal is to provide students with the necessary information; nevertheless, being able to assess whether students have learned is equally critical in order to better meet students' learning requirements. Therefore, assessment and evaluation are essential components of teaching and learning since they are used to measure and gather data about the students' acquired skills and developed language competencies.

Although the terms assessment and evaluation are frequently used interchangeably, they are genuinely two units of the same process: While the assessment is the process of gathering data about what the learner can do, evaluation includes analysis of this collection of data and reaching a decision based on the data (Brown, 2003; Meidasari, 2015). That is to say, "one could look at assessment and evaluation as the journey (assessment) versus the snapshot (evaluation)" (p.224). The data both in assessment and evaluation is yielded

through a systematic approach that draws on multiple sources that can be named traditional and alternative assessment instruments to evaluate the effectiveness of any teaching variable. Consequently, the data gathered through both traditional and alternative tools serves for providing a basis for developing curricula, improving classroom practice and providing insight into the instructional program.

The need for assessing and evaluating students' learning achievement indicates a need to understand the terms evaluation and assessment in more depth. Evaluation is a method that has a pivotal role in determining the effectiveness of the teaching and learning process by knowing whether the materials, the learning strategies, and the assessment procedures are in line with the learners and the curriculum (Gultom, 2016). Accordingly, the tests, observation, or a variety of assessment tools provide a great deal of insight for the evaluation of the teaching and learning process. "Assessment, on the other hand, is an ongoing process that encompasses a wider domain" (Brown, 2003, p.4). It can thus be suggested that assessment is a web that spreads through the teaching process and includes a series of practices that is continually judged by the teacher intentionally or incidentally. That is to say, to do an assessment, a teacher should take many aspects into consideration such as the students' participation, assignments, presentation, performance, or portfolio (Gultom, 2016). Another source of measuring students' performance is a test. In other words, a test is an instrument that measures the learners' ability, knowledge and performance by offering them some kind of result. It is important to note that tests can measure one's ability to perform language in terms of different skills. In addition, one should bear in mind that it is crucial for a well-constructed test to provide an accurate measure of the learner's ability in a given domain (Brown, 2003).

A key issue here can also be the different categorizations of assessment. These categorizations can be instrumental in our understanding of the concept of assessment. A significant distinction on this subject was presented by Brown (2003, p.5-7): In his publication, he discusses informal assessments incorporated in classroom tasks as well as

formal tests that are systematic and occur on a regular basis to demonstrate student achievement. Formative assessment, which focuses on evaluating students in the process and assisting them in continuing their growth, and summative assessment, which focuses on measuring what students have learned, were also discussed.

The assessment at the preparatory department of most higher education institutions is primarily based on the instruments such as written and oral tests, portfolios, or projects to evaluate to what extent students have achieved the instructional goals containing both general objectives and specific objectives concerning basic competencies. Therefore, different types of instruments are administered for students to make up a final score to take the final exam, and the final score is established on the percentage of the varying assessment given throughout the year. The underlying reason for adopting such an approach seems to provide detailed and broader information about students' performance, their present achievement, areas of weakness and strengths, progress, and potential. Therefore, after the coursebook and the supplementary materials are chosen by the department, instruments and methods are decided on accordingly with a view to measure and evaluate whether the students can achieve prescribed objectives and goals.

Considering all of this, it seems that students are one of the most important actors that need to be taken into consideration in terms of instruments that aims to reveal whether a department's instructional goals and objectives are achieved. Variety of issues including whether teaching has been effective and students have learned, or the ways to best address student learning needs necessitate an effective assessment and evaluation process. As this process presents a well-establish link to student performance, it is also a valuable source for developing curriculum and the instructional program. Mediasari (2015, p.224) argues that "the assessment and evaluation literacy from the learner's perspective is also an important part of an instructional program." Thus, it is critical to provide students with the opportunity to express their attitudes towards these instruments which are the essential components of testing.

In view of all that has been mentioned so far, one may suppose that students' attitudes, feelings and opinions about instruments can have a considerable amount of impact on their academic performance. There is a consensus among researchers that emotions have a major role in learners' performances in language learning and the investigation of these emotions, feelings, and opinions would be practical to better understand their impacts on learning (Davidson, 2018; Dewaele & MacIntyre, 2014; Horwitz, 1986, 2010). Hence, it could conceivably be stated that both positive and negative attitudes are primary determinants of language performance. In this vein, learners' anxiety and enjoyment levels that can be categorized as negative and positive feelings emerged as reliable concepts to develop a full picture and generate a well-established insight into the attitudes of learners.

It's important to define the terms "enjoyment" and "anxiety" in this context. Although enjoyment is defined in different fields such as sports and work, the definition of enjoyment as a positive emotion has some common characteristics. A definition for the concept of enjoyment might be that enjoyment is the combination of pleasant feelings, that is, what is considered intrinsically interesting, plausible or satisfactory (Davidson, 2018). Even if the definition may not explain all of the feelings of enjoyment, it can at least provide a useful understanding. As for anxiety, it has been established that anxiety can be defined as a mental and physical state characterized by an unpleasant emotional state or condition (Kráľová, 2016). Just as enjoyment is used in different contexts, anxiety is also used in a variety of situations. However, when it is associated with learning a foreign language, the term anxiety embodies a multitude of concepts which include self-perceptions, beliefs, feelings, and behaviors of learners related to foreign language learning (Horwitz, Horwitz, & Cope, 1986).

There appears to be some agreement that both enjoyment and anxiety are linked to both the cognitive and emotional states of individuals. Moreover, what is notable about these concepts is that both anxiety and enjoyment can be conceived as situation-specific

situations (Huynh, 2021). In particular, this thesis addresses foreign language assessment and evaluation instruments, or for short testing, as a situation-specific situation. Therefore, investigating the anxiety and enjoyment levels of the language learners and revealing their sources of them can provide a useful account of a further and deeper understanding of the assessment and evaluation process.

The investigation of those affective factors in terms of the uniqueness of testing instruments from the perspectives of language learners can make an important contribution to the field. Thus, the present study aims to provide an important opportunity to advance the understanding of foreign language anxiety and enjoyment associated with the assessment and evaluation process. This study, therefore, intends to be of assistance to support existing practices, a guide to discovering new paths, and allow a deeper insight into having a more effective assessment and evaluation program.

Aim and Significance of the Study

The concept of anxiety and enjoyment in language learning is based on language students' descriptions of specific aspects of their courses that are specifically anxiety-provoking or evoke positive emotions (McIntyre, 2017). Thus, it is safe to state that students can also associate language assessment and evaluation tools with some affective factors such as anxiety or enjoyment. It is a widely held view that anxiety and enjoyment are two significant factors in learners' success in language classes (Özer & Altay, 2021). It is therefore likely that revealing students' attitudes towards assessment and evaluation instruments can contribute to the learning environment as learning a language is a complex and multifaceted process involving both learning and being assessed.

Specifically, while researching anxiety and enjoyment as focal elements, it is also in the scope of this study to examine the assessment and evaluation of English as a foreign language, which has long been a question of interest in many classic studies in ELT. What is distinctive about this study is that it indicates students' attitudes in interaction with the

assessment and evaluation instruments. In light of the importance to provide a holistic point of view, this study examines students' attitudes using both qualitative and quantitative lenses to reveal how assessment and evaluation instruments affect their anxiety and enjoyment levels. Therefore, besides giving new insights into the assessment and evaluation of English as a foreign language by involving affective factors, this study makes a major contribution to research on FLA and FLE by offering a complementary view to the issue involving language learners' attitudes on assessment and evaluation instruments.

There have been several studies that focus on language assessment in terms of different skills which aim to improve teachers' assessment practices or institutions' assessment systems specifically based on teachers' views. To date, however, there have been few studies focusing on students' attitudes, and feelings about different types of assessment and evaluation instruments by considering their anxiety and enjoyment levels to provide a broader, deeper, and inside look into the assessment practices. In this sense, Özer and Altay (2021) also emphasized that in a language lesson, learners may experience both anxiety and enjoyment, therefore focusing on a single emotion does not provide the required information to comprehend learners' attitudes and emotional processes during the language learning process.

In this respect, this study attempts to investigate the attitudes of English language learners studying at a university preparatory school towards different assessment and evaluation instruments. The results which are obtained from students' perspectives can contribute to this field by adopting a student-centered approach to provide some insight into the institutions to achieve the stated goals with a student-friendly approach.

Research Questions

Depending on the purposes of this study, the subsequent research questions and sub research questions are formulated;

1. What are the attitudes of A1-level students towards assessment and evaluation instruments?
2. Which of the assessment and evaluation instruments causes the highest and the lowest anxiety among A1-level students?
3. Which of the assessment and evaluation instruments causes the highest and the lowest enjoyment among A1-level students?
4. Are there any significant differences in the attitudes of A1-level students regarding the extent of anxiety and enjoyment in terms of gender?
5. Are there any significant differences in the attitudes of A1-level students regarding the extent of anxiety and enjoyment in terms of academic major?
6. What factors emerge from speaking exams that affect participants' attitudes in the context of anxiety and enjoyment?
7. What factors emerge from quizzes and mid-term exams that affect participants' attitudes in the context of anxiety and enjoyment?
8. What factors emerge from the writing portfolio that affects participants' attitudes in the context of anxiety and enjoyment?
9. What factors emerge from online components that affect participants' attitudes in the context of anxiety and enjoyment?

Assumptions

The main assumptions of this study are as follows:

It is assumed that A1-level students that are studying at a preparatory school are expected to participate in both quantitative and qualitative parts of this research voluntarily. Thus, both quantitative and qualitative data would be collected for this study. In addition, all of the participants in the study would be at the same level of proficiency in the English

language which is A1 level according to The Common European Framework of Reference for Languages (CEFR).

The data collection procedures used in this study (questionnaire and semi-structured interviews) are expected to generate accurate replies from the participants.

The participants in this study are expected to react honestly and candidly to the questions in the data collection instruments (questionnaires and semi-structured interviews).

Limitations

The present study has some limitations. The data for this research was collected online from a state university in Ankara. Participants were encouraged to engage in this research by filling out online forms willingly because the data collecting period corresponded with the COVID-19 pandemic and online education was a viable alternative for the university where the study was conducted. In addition, the researcher used WEBEX to interview the students (an online video chat program). Therefore, the major limitation of this study is the absence of opportunities to gather data face to face.

This study's data collection period is limited to the academic years 2021-2022. It's possible that extending the data gathering duration will result in better findings. As a result, another drawback of this study is the data gathering time.

The item pool of the scale was created using a comprehensive literature evaluation in this study. After the item pool was created, a five-point Likert questionnaire was created, and expert opinion and piloting were conducted. The number of students who participated in this study is another drawback of this study due to the difficulties of reaching every university student in Turkey.

The examination of students' attitudes regarding assessment and evaluation instruments is limited to a few essential elements in this study. The last limitation of this research is that the assessment and evaluation instruments are those used in online

education due to the Covid-19 Pandemic. As a result, the participants' attitudes are limited to online assessment and evaluation instruments administered at a state university in Ankara.

Despite the aforementioned limitations, this study is important in the study of attitudes toward various assessment and evaluation instruments in the context of foreign language learning.

Definitions

Testing: It is a method of evaluating the progress of students and determining student outcomes and individual student needs (Overton, 2012)

Assessment: The process of obtaining data in order to track development and, if necessary, make educational decisions. It encompasses a variety of styles and methodologies, including formative vs. summative, informal vs. formal, continuous vs. final, and process vs. product evaluation. (Harmer, 2007; Overton, 2012).

Evaluation: Procedures used to determine whether the subject (i.e. student) meets preset criteria. This uses assessment such as a test to decide on qualification in accordance with predetermined criteria (Overton, 2012).

Assessment and evaluation instruments: Practices that can include both traditional instruments such as quizzes and tests and alternative and innovative practices such as portfolios, speaking tasks, or projects (Armstrong, Marine & Patarrayo, 2004).

Language anxiety: Language anxiety, often known as foreign language anxiety (FLA), is a multi-faceted term that is classified as situation-specific anxiety that might obstruct the acquisition and/or production of a second language (Horwitz, 2010).

Enjoyment of language learning: It is a pleasant feeling that can be defined as an experience of fulfillment and joy that occurs when challenges and skills to fulfill them are well aligned (Dewaele & MacIntyre, 2014).

Chapter 2

Theoretical Basis of Research and Literature Review

The conceptual framework and prior studies on the relevant literature are presented in this chapter. In the first section, the main units of testing and the importance of assessment and evaluation in foreign language learning are touched upon. Following that, positive and negative emotions in language learning related to attitudes, foreign language anxiety, and foreign language enjoyment are covered in detail. In the last section of this chapter, previous studies conducted on all these subjects are mentioned in their relevant sections.

Language Assessment and Testing in General

Assessment and testing that is significant content knowledge areas are the fundamental components of teaching and learning activities (Allen, 2005) and they can be traced back considerably further in time with various patterns being pertinent to the academic field (Spolsky, 2016). Multiple sources of information are used in the assessment that is an integral part of the teaching process to increase the chances of better interpreting what students have learned. It is therefore likely that clarifying the important terms - "assessment, testing, and evaluation" - seems among the essential things in achieving the aforementioned goal.

In the literature, the terms "assessment", "testing", and "evaluation" have been frequently used interchangeably to refer to various types of judgments or conclusions concerning language learning and its components (Brown & Hudson, 1998). They do, however, vary from one another to some degree. Assessment is a way of gathering information about a learner's skill and knowledge in order to determine whether or not the learner can meet the objectives (Brown, 2003). That is, assessment is an umbrella phrase that encompasses a wide range of tools for evaluating learning using obtained data in order to improve learning. Therefore, Lynch (2001) stated language tests are included in the

assessment, which is a broader term. In this regard, assessment and testing are not exactly interchangeable and it can be suggested that the former encompasses the latter. Since assessment encompasses testing, it can also provide teachers with both qualitative feedback and quantitative information through tests and examinations as a useful measure of the learning process to tailor their teaching practice (Gökhan, 2004; Tosuncuoğlu, 2018).

To date, language testing has evolved as a result of significant new advancements, and it has faced various obstacles and complications as a result. Researchers have used numerous interdisciplinary techniques to address old and new challenges in language learning and evaluation during the last few decades, and the area has gradually evolved in scope and sophistication (Aryadoust, Eckes & In'nami, 2021). In other words, more has been emphasized along with the assessment of the four language abilities (listening, speaking, reading, and writing) engaging with the idea that assessment and learning are inextricably linked. (Aryadoust, Eckes & In'nami, 2021).

Testing

It is explicitly stated that testing is an indispensable tool in today's language learning environment, and that it is the flipside of teaching (Tamunobeleva, 2016). Testing is one way to assess progress and determine student outcomes as well as particular student needs. It is simply one form of assessment (Overton, 2012), and therefore it refers to a method for determining a student's capacity to accomplish certain activities and demonstrate mastery of a skill. Testing has a long history, and many educationalists prefer it to evaluate student success and proficiency because exams appear to be more reliable in their eyes with respect to the fact that “they have the ability to observe, perform surveillance, quantify, classify, normalize and judge” (Shohamy, 1993, p.2). As a result, testing has become a common aspect of the teaching and learning process.

Tests are typically made up of a high number of questions and items, each of which includes a problem presentation and a range of alternative answers (Berlak, 1992). An exam will always have a time restriction, i.e., there is always specific time devoted to exams, and

results are normally calculated by counting correct responses (Berlak, 1992; Shohamy, 1993). Some tests may feature open-ended test items that necessitate the submission of a writing sample and in scoring such items, a person who is also skilled in scoring assigns a number to each response (Berlak, 1992). Ultimately, aside from the type of test, learners must be tested to see how far they have progressed in any area of language, as well as data for various uses by education stakeholders (Tamunobelega, 2016).

The Types of Language Tests. Teachers are expected to assess student competency and understanding of educational goals and objectives using a range of assessment strategies (Hughes, 1989; Overton, 2012). In this context, many sorts of assessments can be utilized to obtain objective evidence for evaluating the program's efficacy and determining which areas require additional exploration. To better understand the types of the tests, Hughes (1989, p.9-14) classified them depending on the purposes of the tests:

- Language proficiency tests are used to evaluate a person's ability to communicate in a foreign language.
- Language accomplishment tests are used to determine whether or not students have met the course goals.
- Diagnosis tests are used to identify a learner's flaws and strengths, as well as the elements they already know and don't know.
- Placement tests are used to place students in the most appropriate instructional program for their abilities.

There are also some different approaches to test construction:

- Norm-referenced tests are those that evaluate individual students to national averages or norms of expected outcomes (Overton, 2012). On the other hand, Criterion-referenced tests categorize pupils based on their ability to do certain tasks satisfactorily; they can pass if they meet the standards and perform well;

those who do not will fail (Hughes, 1989). This means that evaluating their achievement is based on significant criteria rather than on the performance of their peers.

- Testing is said to be direct when it demands the learner to perform precisely the skill we want to evaluate, whereas indirect testing attempts to measure the abilities that underpin the skills we want to measure (Hughes, 1989). The presence of a single language element, or a combination of language elements, appears to be significant in this case. As a result, a test can evaluate one piece at a time or require the learner to integrate multiple language items in a task. The former is called discrete point testing while the latter is called integrative testing (Hughes, 1989).

The various approaches and test types have been discussed in this section. Therefore, it seems that the need for a better interpretation of learners' performance and a more accurate measurement of the abilities that educators are interested in has prompted the development of a variety of approaches and test types aimed at providing testers and teachers with the fundamental knowledge and techniques to improve testing.

The Impacts of Tests. Originally intended to offer information on achievement or to identify suitable candidates, tests have evolved into tools for regulating educational systems (Cumming, 2009; Shohamy, 1993). As we need to know where we are today in order to know where to go next in any process, testing and assessment seem the sole ways to verify students' language improvement (Cumming, 2009; Tomunobeleva, n.d.). Tests have social, psychological, curricular, and educational consequences, and administrators, teachers, and students are the ones affected by these results (Shahomy, 1992).

Although tests are typically used to drive and promote language learning, they may have both positive and negative repercussions for stakeholders. The impact of testing on teachers and students cannot be overstated, particularly in institutions where language tests are the main form of student assessment. Since it is possible that students are not

presenting a representative performance of their language competency if the test has flaws, how students go through the process of taking language tests might affect the learners' learning process (Cohen, 1984). If the test is not valid or does not reflect genuine teaching and learning, it can tarnish teachers' reputations and cause conflict between administrators and teachers (Shohamy, 1992). It could even be a lack of awareness on the part of teachers of what might be an adequate strategy to prepare students for the examination, aside from the test itself (Alderson & Wall, 1993). As a result, not only between students and teachers but also between testers/test developers and teachers, this can have negative consequences. It can thus be suggested that the impacts of tests encompass the attitudes of individuals and groups including teachers, students, administrators, testers and even publishers regarding the test.

Washback Effect. A variety of terms highlight the relationship between testing and learning: The washback effect is one of them; it refers to the impact of testing on teaching and learning (Shohamy, 1993). It is also defined as the effect of testing and assessment on the language education curriculum by Brown and Hudson (1998). Although there is broad agreement in the area on the basic description of washback, there is also a great deal of variation on how it works. Although washback can have a good or negative effect, it is usually believed that particularly negative washback exists (Bailey, 1996). Based on Brown and Hudson's definition of washback, this condition could be linked to a negative washback effect if the curricular objectives and what is tested are not in accord with one other. In contrast, if the test is successful in assessing what is outlined in the objectives and what is delivered in the classroom, a positive washback impact can be stated. It has been conclusively shown that washback is described as the impact of testing on teaching and learning, and it is commonly accepted that it exists and is significant (Bailey, 1996).

The concept of 'washback,' as implied by the definitions, is widespread, and there are claims for both negative and positive washback. Negative washback is attributed to a lack of test validity by some writers, while others attribute it to other factors such as test

anxiety or test type (Alderson & Wall, 1993). As a result, washback is a multifaceted issue that can be induced by a variety of circumstances; however, it should be emphasized that the washback effect is one of the impacts of tests, which refers to a larger concept defined above. Furthermore, exams that are administered under a time constraint and thus may put pressure on the students may be insufficient to indicate the students' actual and holistic performance and therefore should be reinforced by other assessment instruments. As a result, washback, a term primarily associated with testing, is already a complex matter and examining it appears to be a challenging endeavor.

Assessment and Evaluation

Evaluation and assessment are critical in the development of an institution as well as an individual's performance or success. The process of systematically judging or determining one's abilities or performance is defined as evaluation (Rao, 2018). What is notable about this definition is that it actually stresses that evaluation is an integral part of the teaching-learning process since it gives useful input on the program's design and delivery and enables both teachers and students to develop their abilities on a regular basis.

Assessment is the process of obtaining data in order to track development, detect learners' weaknesses and, if necessary, make educational decisions; therefore, many formal and informal techniques for evaluating student development and conduct are included in the assessment (Overton, 2012). The word 'assess' comes from the Latin verb 'assidere,' which means 'to sit with.' : It implies that teachers work with students to help them improve their performance through continual assessment (Rao, 2018). It is necessary here to clarify exactly what is meant by his statement: The teacher examines the student's replies and assesses his or her ability to perform a certain task. If a student makes a mistake, the teacher determines the sort of mistake and the procedures that must be followed to fix it. The teacher observes the student's behavior, accumulates information about the student, and adjusts the instructional strategy accordingly (Overton, 2012). As a result, it can be

suggested that the term "assessment" refers to a procedure for identifying strengths as well as weaknesses and, as a result, heading toward a problem-solving process.

Types of Assessment. Educational institutions, as well as the teachers that work there, may maintain or follow particular methods or standards in order to assess student performance. Therefore, the assessment that is done in the English language teaching-learning process is classified into different types:

Summative Assessment and Formative Assessment. Teachers are frequently asked to not only evaluate their pupils' knowledge but also to summarize that evaluation into a letter or number grade so that they can see if they are gaining this academic knowledge (Allen, 2005). This is known as "summative" evaluation. In other words, a summative assessment is one that is done at the end of a period of education to identify the amount of skill development or mastery (Overton, 2012). On the other hand, formative assessment is a type of continuous evaluation that takes place as a person learns a new ability. Therefore, "formative" assessment is a type of continuous evaluation that takes place as a person learns a new ability (Overton, 2012) and some questions can be directed for formative assessment (Gultom, 2016, p.191):

- Has enough time been spent on certain goals?
- What kind of response has the textbook gotten?
- Is the methodology used by teachers appropriate?
- Is the program a hit with the students? If not, what can be done to help them become more motivated?
- Is the material's pacing appropriate?

As is seen, the basic goal of formative assessment is to provide continuous feedback to both teachers and students (Rao, 2018). Teachers should also collect nongraded "formative" evaluations of students in order to provide feedback to them as they learn, as well as consider how to motivate students to learn and help them in becoming self-regulated learners. (Allen, 2005).

Informal Assessment and Formal Assessment. Assessment is an important part of in-class activities, and a teacher can respond informally and formally to a student's performance in order to interpret information about the learner in the most objective way possible. As Brown (2003) and Overton (2012) suggested that teachers might use informal assessment tools to track their pupils' academic progress by checklists and questionnaires, as well as reviewing student work samples. In other words, informal assessment is a way of providing feedback and comments on students' performance in the classroom which is frequently unplanned and spontaneous. Formal assessments, on the other hand, are structured, planned, and prepared beforehand to be administered to determine what skills and content the student has mastered (Brown, 2003). In this sense, tests are well-known as vital methods for formally assessing learners' performance.

Objective Assessment and Subjective Assessment. The purpose of objective and subjective assessments is to draw conclusions about learners' linguistic proficiency. In his book, Carl (2001) identifies objective and subjective assessments based on how to score a test: A test that can be graded objectively is known as an objective test such as multiple-choice questions, as well as true-false or matching questions. A subjective test, on the other hand, is one that is scored using human judgment, such as most writing or speaking tests. Teachers' grading of writing tasks is an example of subjective evaluation in this sense, as the grading is mostly based on the teacher's subjective appraisal of the relevant performance. There are clearly some methods for overcoming the drawbacks of subjective evaluation. More than one scorer is preferable for increasing assessment reliability since it improves objectivity. As Hughes (1989) stated that the more consensus between scorers, the more objective the scoring becomes. A multiple-choice item or a true-false question, on the other hand, does not necessitate the scorer's judgment.

Discrete Point Assessment and Integrative Assessment. Discrete-point assessment is built on the premise that a language can be broken down into its constituent pieces and that those parts can be properly tested (Brown, 2003). To put it another way,

various skills such as listening, speaking, writing, and reading, as well as language units such as lexicon, syntax, and morphology, are tested separately. On the other hand, cloze tests and dictations are great examples of integrative assessment since they aim to measure overall ability by stressing the interconnectedness of different abilities and points of language (Brown, 2003; Carr, 2011). Integrative assessment, in other words, suggests an indivisible view of linguistic proficiency by requiring test-takers to blend diverse features of a language into the task.

Norm-referenced Assessment and Criterion-referenced Assessment. A norm-referenced assessment allows teachers to compare a student's performance to that of other learners of the same age or grade level (Overton, 2012). In this type of assessment, a sample of learners is chosen as the norm group since it is not practical to assess the performance of every pupil of the same age or grade level. Contrary to norm-referenced assessment, the distribution of students' results throughout a continuum may indeed be unimportant in criterion-referenced assessment as long as the instrument assesses relevant objectives (Brown, 2003). The difference between these two types of assessment is that norm-referenced assessment tells test users how well a particular examinee performed in comparison to other examinees, whereas criterion-referenced assessment can reveal how well that individual performed in absolute terms (Carr, 2011).

To summarize, assessment can take many different forms depending on the assessment objectives. As formerly stated, each of these types of assessments has its unique set of characteristics, which must be considered while selecting and administering them.

Assessment of Language Skills

For decades, applied linguists have been debating whether language ability is a single, unitary trait or if it may be divided into separate components (Powers, 2010). Listening and reading are regarded to be receptive skills, while speaking and writing are productive skills. As a result, the four fundamental skills are linked since students are

expected to use language for different purposes such as producing oral or written work and receiving and producing messages (Harmer, 2007; Powers, 2010). Therefore, it seems that capturing a detailed and comprehensive depiction of a test taker's overall proficiency in four English skills is crucial.

Assessing Reading Skill. Teachers prioritize getting students to read English texts for a variety of reasons: Reading improves students' vocabulary, spelling, and writing skills, because reading texts provide good examples of vocabulary, punctuation, and grammar (Harmer, 2007). Thus, improving reading skills while learning a language is an essential and inextricable aspect of the process.

For assessment purposes, language learners are expected to do several things with a reading text. Burgess and Head (2005) identified the following criteria for assessing reading: Candidates are tested on their ability to use skills such as skimming, scanning, inferencing, detecting discourse markers as well as recognizing grammatical word classes, patterns, and coherent devices. They also emphasized that the length of the text, the subject matter, the style, and the language employed are all crucial factors to consider while assessing reading. Furthermore, given that students have limited time in a reading exam, it is reasonable to simplify the tasks and texts in order to present texts at a level of difficulty that is appropriate for the students (Brown, 2003).

Assessing Listening Skill. Most language learners want to understand what others are saying to them in the target language, thus it is important for them to be exposed to a variety of listening sources. In this sense, listening texts are crucial because they serve as strong pronunciation examples, allowing students to improve not only their knowledge of speech but also their ability to speak (Harmer, 2007).

Candidates' listening skills are typically assessed by having them listen to audio recordings and complete written assignments to demonstrate their comprehension of the spoken language (Burgess and Head, 2005). Therefore, listening assessment can be defined as a test of a learner's ability to process spoken language. In listening texts, a

variety of themes, topics, and task types are available, so a tester first determines the aim of the assessment and builds the test accordingly. Brown (2003) distinguished four categories of listening performance: Intensive, extensive, responsive, and selective listening. To begin with, there is a distinction to be made between intensive and extensive listening. The latter refers to listening that students perform outside of the classroom for pleasure or other purposes, whereas the former refers to listening that is done purposefully to improve listening skills and study how English is spoken (Harmer, 2007). Students are expected to listen and reply in responsive listening, therefore their performance can be measured by composing a response, selecting an appropriate response, and providing an oral response (Brown, 2003). Students are needed to listen and scan for specific information through assessment techniques such as information transfer and phrase repetition when it comes to selective listening (Brown, 2003).

Hughes (1989) and Harmer (2007) emphasized that when it comes to testing preparation, test specifications should be carefully considered while selecting listening texts, genres, and activities as well as redundancy, authenticity, and the quality of the recordings. To summarize, teachers play a critical role in ensuring that students can use a variety of listening abilities to grasp the overall meaning or specific details, and both teachers and testers possess the essential knowledge to select and design listening tests.

Assessing Writing Skill. For a long time, writing seems to be regarded as only a tool for learning grammar and vocabulary, rather than a skill in and of itself. After a while, the value of teaching and assessing writing in foreign-language classrooms, on the other hand, has been acknowledged and highlighted (Harmer, 2004).

Brown (2003) attempted to draw fine distinctions concerning the types of writing: Imitative, intensive, responsive, and extensive writing. To begin with, imitative writing is used to examine students' understanding of language mechanics such as spelling. This sort of writing frequently employs spelling tasks, picture-cued, and matching phonetic symbols, therefore; it's a stage where students are attempting to master the mechanics of writing. In

intense writing, learners are needed to write words and phrases utilizing the right structures within a context; consequently, meaning and context play a role; yet, there is more focus on form. As for responsive writing, the main focus is on meaning and context because learners are needed to execute at a limited discourse level, connecting sentences into paragraphs and generating a logically connected sequence of two or three paragraphs. Lastly, extensive writing entails mastering all writing processes and strategies for all purposes; as a result, students are required to attain a goal and arrange and develop ideas logically while displaying syntactic and lexical variety (Brown, 2003).

When choosing between the types to design the writing task, teachers and testers should keep their assessment purpose in mind. Aside from that, there are a few more considerations to take into account. Hughes (1989) highlighted some important aspects to consider: The task should solely assess the students' writing ability, not their background knowledge. Furthermore, providing task requirements to students with basic and clear directions is usually useful.

There are a number of ways to assess, respond, correct writing tasks, and guide students. Teachers are asked to play a variety of roles when assessing and providing feedback on students' written work, as well as the opportunity of using alternative scoring methods. Hughes (1989) presented two methods of scoring: Holistic and analytic: The focus of holistic scoring is in the scorer's overall opinion of the task, whereas analytical scoring demands scorers to assign separate points to pre-defined areas of writing skill. In his book, Harmer (2004) mentioned different ways of correcting students' work: A teacher can adopt selective correction that focuses on correcting only pre-determined/pre-discussed parts of language such as verb tenses or word order. S/He can use marking scales by giving marks out of 10 for each category (e.g. grammar, vocabulary, coherence, or cohesion). Additionally, using correction symbols that encourage students to consider what they've done wrong, reformulation that allows students to compare correct and incorrect versions, and referring students to a dictionary or grammar book to encourage them to look at

information with the goal of correcting their errors are some other ways mentioned in Harmer's (2004) book.

To summarize, assessing writing ability is not a simple and direct task for teachers; consequently, language teachers need to understand the various types of writing and the skills required to assess students' writing abilities in order to properly assess the skill.

Assessing Speaking Skill. There are three key factors for getting students to speak, according to Harmer (2004): First, students can practice speaking in real-life situations in the safety of the classroom. Second, owing to the teacher's feedback, everyone can see how successful they are. Finally, speaking allows students to activate various parts of language stored in their minds. It can therefore be stated that speaking is a productive skill that can be observed and assessed using specific criteria.

Oral performance can be divided into five categories (Brown, 2003; Harmer, 2004):

- Imitative kind of speaking requires learners to repeat a word, phrase, or sentence to be able to check pronunciation.
- Intensive speaking necessitates learners to produce sentences with little or no contact through picture-cue tasks or dialogue completion.
- Responsive speaking involves some interaction and comprehension.
- Interactive speaking necessitates a more complicated and prolonged connection through conversations, role-plays, and interviews.
- Extensive speaking necessitates the development of increasingly complicated oral skills, such as giving presentations or speeches.

As for scoring, the speaking ability can be scored using both holistic and analytic methods by using a scale; nevertheless, it is critical to have more than one scorer, train them, and evaluate the multiple parts of speaking skill at the same time (Hughes, 1989). In short, evaluating a productive language skill necessitates meticulous planning. To put it another way, it's critical to construct speaking exams based on the assessment's goal in order to accurately measure students' speaking abilities.

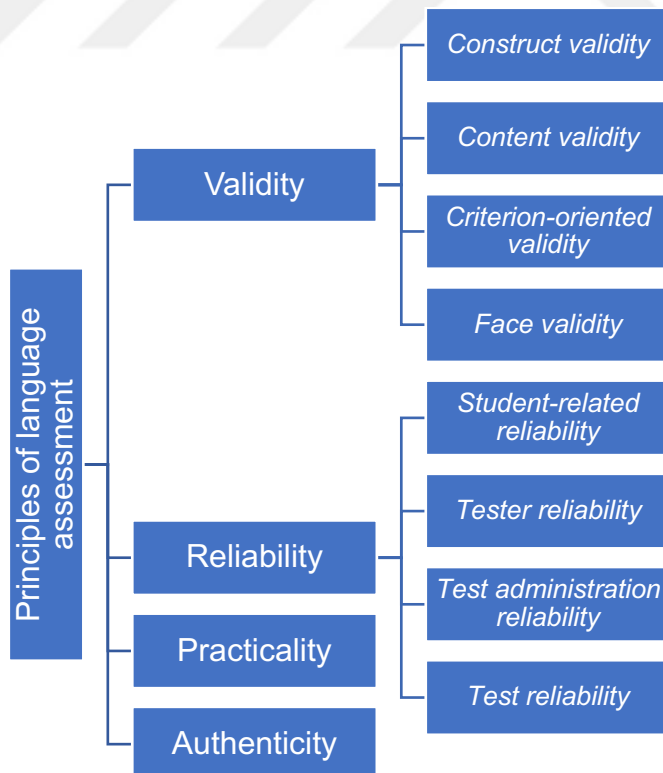
Principles of Language Assessment

Test development and assessment processes are cyclical or continuous in nature, as they may require going back and making adjustments or rewriting. This is due to the fact that there are numerous principles to consider, and the goal is to construct a test or assessment instrument that is as effective as feasible for the purpose intended. Furthermore, understanding reliability and validity will help educators decide when to utilize a particular instrument (Overton, 2012).

To create a good test and conduct a successful assessment process, one must first ask how we can tell if a test is useful, and the answer is that there are four criteria for testing a test: Validity, Reliability, Practicality, Authenticity (Brown, 2003). Each principle is described in the following section, along with any sub-branches that could exist.

Figure 1

Principles of Language Assessment



Validity. The validity principle, possibly the most important, is concerned with measuring ability in a relevant and suitable manner. That is, if the reading ability is to be assessed, a valid test measures reading rather than prior knowledge of a subject or some other variable of debatable relevance (Brown, 2003; Carr, 2011). As Hughes (1989) and Overton (2011) also stated that a test is said to be valid if it measures what it is designed to assess. In other words, validity concerns whether a test measures the things it purports to assess. It can thus be suggested that validity is one of the most significant components of successful assessment since it influences assessment quality. There are also several kinds that focus on certain aspects of validity: Construct validity, content validity, criterion-oriented validity, and face validity are the four types of validity described below.

Construct Validity. The degree to which a test score can be interpreted as indicative of the construct of interest is known as construct validity which is seen as the essence of validity (Carr, 2011). Brown (2003) stated that constructs refer to hypotheses or models that attempt to observe a phenomenon that can be measured directly or experimentally. In this regard, proficiency and communicative competence can be thought of as linguistic constructs, and construct validity is concerned with determining if the test genuinely taps into the theoretical construct as defined (Brown, 2003). This can be illustrated briefly by an example. For instance, certain variables like content, cohesive devices, language (suitable, accurate, and varied grammar structures), mechanics, and vocabulary are taken into account when grading a written task since the rationale for these factors is based on a theoretical construct. When we evaluate a written task based on simply one or two of these variables, however, it indicates that the assessment lacks construct validity.

Content Validity. When the items in a test are indicative of the content that is supposed to be measured, it is said to have content validity (Overton, 2011). A further definition is given by Carr (2011) who describes content validity as the adequacy with which the test items sampled the "universe" (i.e. domain) to which the test developer intended

scores to generalize. In other words, content validity requires that a test contain the content in a representative manner. In a situation where a teacher utilizes a reading test that only evaluates one aspect of reading such as word recognition and uses the obtained score to place a student, believing that the student will be able to comprehend reading material at a particular level, this poses a problem for content validity (Overton, 2011). It is apparent that the test's content validity for assessing overall reading ability is insufficient. That is to say, if you can clearly define the achievement you're assessing, you can obtain content-related evidence through observation (Brown, 2003). To assess a student's speaking ability, a teacher must provide a task where a student can actually speak in order for the assessment to be content valid.

Criterion-oriented Validity. The degree to which the findings of a test correlate with the results of another independent and dependable assessment that the examinee participated in for a given ability is referred to as criterion-oriented validity (Brown, 2003). According to a definition provided by Carr (2011) and Overton (2011); while the degree of correlation between test scores and scores on another test that was designed to assess the same construct is known as concurrent validity, the relationship between test scores and later achievement on anything is known as predictive validity. Consequently, concurrent and predictive validity create criterion-related validity.

Face Validity. When a test appears to measure what it claims to measure, it is said to have face validity, which is scarcely a scientific term (Brown, 2003). Thomas et al. (1992) use the term face validity as a multifaceted, complicated term that can be used to evaluate how test items seem to responders and others. In this sense, it can be assumed that a test is said to have face validity when a test-taker believes it is valid and meets their expectations. On the other hand, face validity would be considered lacking in a test that professed to measure pronunciation abilities but did not require a candidate to speak (Brown, 2003).

To sum up, it can be stated that while content and face validities play a prominent role in the design of a test, construct and criterion validities have come to the forefront in the pre-application phase (Yurdugül & Bayrak, 2012).

Reliability. What the literature suggests is that grading is arguably the most essential measuring decision, so the concept of reliability is also one of the most important assessment concepts (Allen, 2005). The consistency and accountability of a test are referred to as reliability, which describes a condition in which the test provides similar findings at different administration periods. (Brown, 2003; Hughes, 2003). In this respect, a test is reliable in the sense that it measures a trait with roughly the same degree of accuracy throughout time (Overton, 2012). A prominent example of what is meant by reliability is that if a student gets substantially different scores when taking the same test in a short period of time, there has to be a problem with the test's reliability since the student's English proficiency level cannot change in that time. Brown (2003) identifies four aspects that affect reliability: Student-related reliability, rater reliability, test administration reliability, and test reliability.

Student-related Reliability. Students' physical or psychological states on test day, such as weariness, demotivation, and anxiety, are student-related factors that affect the test's reliability (Brown, 2003).

Tester Reliability. As previously stated, a candidate cannot obtain the same exact score on two examinations conducted on different days. In this situation, we may also suppose that one scorer will not provide the exact same score both times (Hughes, 1989). However, reaching a degree of agreement in the ratings of the same scorer on these two instances, as well as between the scores of two distinct raters, seems likely and necessary. Inter-rater reliability is low when there is a discrepancy in the scores of different raters. Intra-rater reliability, on the other hand, falls when the scores given by the same rater are inconsistent. (Brown, 2003).

Test Administration Reliability. Brown (2003) addresses that physical circumstances such as lightning, noise, temperature, copy quality, and the seating arrangement in which examinations are delivered can all affect reliability. The exam's reliability will be compromised if students are bothered by inadequate lights and a low-quality test copy.

Test Reliability. Tests that are too long or too short, as well as tests with confusing instructions, are problematic in design, which means they are unreliable because test-takers may feel overwhelmed by the amount of time spent or by the amount of uncertainty created by the test (Brown, 2003). Therefore, it's also possible that the tests themselves have an impact on reliability.

Practicality. Another element of an effective assessment that should be considered while administering a test is its practicality. A practical test is one that is effective. This means it isn't very expensive, remains within reasonable time limitations, is generally simple to administer, and uses a time-saving scoring procedure (Brown, 2003). In this sense, it can be conclusively suggested that time, location, staff, supplies, and equipment are all resources that must be considered when conducting a test (Carr, 2011).

Authenticity. When you claim authenticity in a test task, you're suggesting that the task is likely to be performed in the actual world, which means (Brown, 2003):

- The test language is as natural as feasible.
- Rather than being separated, items are contextualized.
- The learner can relate to the topics.
- Real-life tasks are represented by tasks.

To put it another way, the exam tasks cannot be separated from real-world tasks and isolated from the learner's real-life experiences to draw judgments about what a test taker can do with the language in the real world (Carr, 2011). To summarize, the principles listed above are critical for designing and evaluating tests. It is critical that language

teachers, test developers, and administrators consider these ideas when evaluating students' abilities.

The Attitude Concept

As a research object in social psychology since the 1900s, and as a current research object in numerous domains such as education sciences, economy, and politics, there are many definitions of the attitude concept (Akdamar, 2020). In everyday English, the word attitude has a wide range of meanings, and it took decades for the concept to crystallize into any type of agreed-upon definition (Reid, 2015).

An acquired tendency to react positively or negatively to a specific object, circumstance, organization, concept, or other person is known as attitude (Tezbaşaran, 2008). Opinions and attitudes differ from each other, yet they are comparable in terms of generalizability and measurement technique. While opinions emphasize precise reactions to certain events or situations, attitudes include a person's reactions to a greater collection of events or groups of individuals (Tezbaşaran, 2008). Although attitudes cannot be readily seen or observed, they do underpin many elements of behavior, which is why studying attitudes is so essential (Reid, 2015). Therefore, it can be asserted that many are conscious of their own opinions, but not always of their attitudes. To uncover attitudes, first, it is agreed on the attitude/s to measure, and then some statements about the attitude object in relation to the individual's experiences (i.e. thoughts, feelings, and behaviors) are constructed. Thus, attitudes are made up of a small number of beliefs that perform a critical function in making meaning of certain concepts, objects, or circumstances (Reid, 2015). To summarize, attitudes are certainly present, even if they are rarely observed and judged explicitly, and they are typically considered precursors to conscious acts (Swamy, 2007).

Positive and Negative Attitude

There are three ways in which attitude manifests itself (Reid, 2015):

- Cognitive: Knowledge of an object (The beliefs and ideas component)

- Affective: a reaction to something (Like or dislike component)
- Behavioral: a tendency for doing something (how we behave toward the object)

As seen above, attitudes are composed of three parts, and it's vital to understand that attitudes can be directed at someone or something. By considering these components of attitudes in the educational setting, it is possible to argue that a student in a language class may have positive or negative thoughts and feelings regarding certain parts of the language. "Certain parts" are purposely highlighted here because while a student may enjoy in-class speaking activities and find them favorable, he or she may have a negative attitude toward the speaking exam. Therefore, both positive and negative attitudes about a particular language are possible (Youssef, 2012). Some students may have a negative attitude toward the second language while others may have a positive attitude. Indeed, a student can have a positive attitude toward the course, but s/he can have a negative attitude toward the course teacher.

All in all, attitudes about language itself and different aspects of language are an important issue because they affect how students learn languages (Youssef, 2012). Identifying whether students' attitudes toward various aspects of the language learning process are positive or negative can assist teachers in becoming more aware of the situation from the perspective of students, and acting accordingly by using some strategies or making some changes and applying more useful strategies for the implied aspect.

Attitude and Language Learning

Learning a foreign language is a delicate and personalized process influenced by a variety of elements and variables starting with the learner and the teacher. Intelligence, learning style, personality, anxiety, positive, and negative attitudes, age and even gender all can play a role in the learning process. That is to say, there are various elements that influence the learning of a foreign language, and several studies have shown that both cognitive and affective factors influence language learning. Attitudes are actors that are

capable of both aiding and complicating the acquisition of a second language (Krashen, 1982). Based on what he argued, it can be stated that positive attitudes can aid in language learning, whereas negative attitudes can hamper it. From this perspective, it is critical to consider students' attitudes about the target language in order to fulfill instructional objectives.

Nath et al. (2017) put forwards that the affective filter (i.e. motivation, anxiety, and self-confidence) is a mental filter that hinders language learners from being receptive to meaningful language input, disturbing their acquisition process. Gardner (1985) also stated that positive attitudes toward target language acquisition foster motivation, which is essential for foreign language learning success. As a result, adopting a positive attitude toward the target language might help students become more motivated and ready to learn it. In other words, affective elements like negative attitudes, a lack of motivation, anxiety, or positive attitudes including enjoyment, having self-confidence have the ability to either help or hinder language learning.

It appears to be the case that students may have positive or negative attitudes regarding what they are learning, how they are learning, and how they are tested can have some significant effects on the performance of language learners. Reid (2015) suggested that attitudes are made up of aspects of knowledge and have a significant impact on behavior, so people can also use them to make sense of their surroundings, themselves, and their relationships with specific situations, things, or concepts. Therefore, teachers, testers, and administrators may get fresh insights into the language learning process and its essential components by analyzing learners' attitudes, especially in a situation where learners are generally passive in decision-making processes. Furthermore, by studying what attitudes students have, how anxious they are, or how much they enjoy the language learning journey, it is possible to derive an approximation of the efficiency of the program, assessment, and evaluation instruments.

Anxiety in Language Learning

Despite a wide and expanding literature, anxiety is still a very diverse and frequently conflicting idea, so basic definitions, methods, and consequences are all areas where researchers disagree (Sarason, 1984).

The hypothesis that anxiety affects language learning has long piqued the interest of academics, language teachers, and language learners (Horwitz, 2010). People with anxiety are possible to be concerned about the challenges they may face in the future. Thus, anxiety is also likely to appear as an obstacle to learning a second language. Low self in anxious people is not only unpleasant to experience, but they also have negative performance consequences since they are self-preoccupying and hinder task concentration (Sarason, 1984). However, as there is affective involvement, non-affective elements, and potentially confusing outcomes, research into the relationship of anxiety to foreign language learning is neither easy nor well-understood (Scovel, 1978).

Types of Anxiety

Anxiety is a complex notion, and psychologists have classified it into several categories, including facilitative-debilitative anxiety, trait anxiety, state anxiety, and situational anxiety (Horwitz, 2010; Horwitz, Horwitz & Cope, 1986).

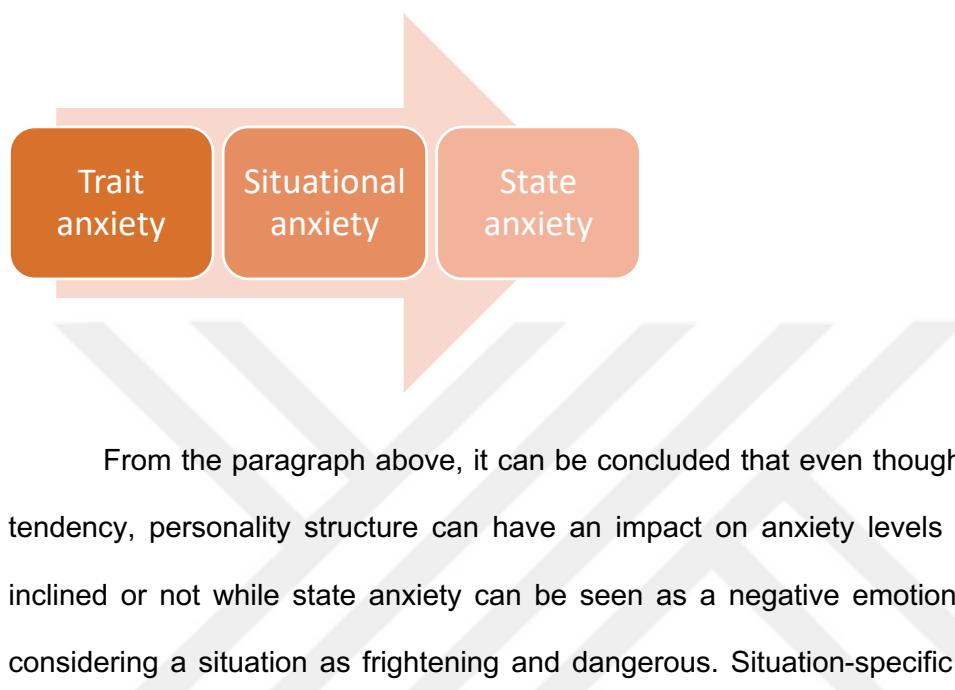
To better understand anxiety, Scovel (1978) classified it into two distinctive types: One of them is facilitating anxiety which can motivate the learner to overcome difficulties to feel emotionally approved and the other type is debilitating anxiety which can discourage the learner from dealing with a new task and lead avoidance.

As for the other three categories mentioned above, Zheng (2008) assumed that the distinctions between them are loosely recognized on a continuum from stability to impermanence; however, there is no obvious division between them: Trait anxiety is characterized by a consistent tendency to be nervous in a variety of settings, state anxiety is linked to the experience of a transitory emotional state from moment to moment and

situational anxiety is in the center of the spectrum, indicating the likelihood of being anxious in a certain situation.

Figure 2

The Nature of Anxiety



From the paragraph above, it can be concluded that even though trait anxiety is a tendency, personality structure can have an impact on anxiety levels as well as being inclined or not while state anxiety can be seen as a negative emotion that arises from considering a situation as frightening and dangerous. Situation-specific anxiety refers to anxiety that is specific to a single context. In this sense, situation-specific anxiety can include language anxiety and such anxiety can negatively affect the language learning process (MacIntyre & Gardner, 1991).

Foreign Language Anxiety

When anxiety is limited to the context of foreign language learning, it is classified as a specific anxiety reaction (Horwitz, Horwitz & Cope, 1986). Horwitz et al. (1986) put forward that language anxiety is a unique combination of self-perceptions, beliefs, feelings, and behaviors associated with classroom language learning emerging from the distinctiveness of the language learning process. In this sense, language anxiety can be viewed in terms of negative thinking and self-related cognition that is undesirable in language classes (MacIntyre & Gardner, 1991).

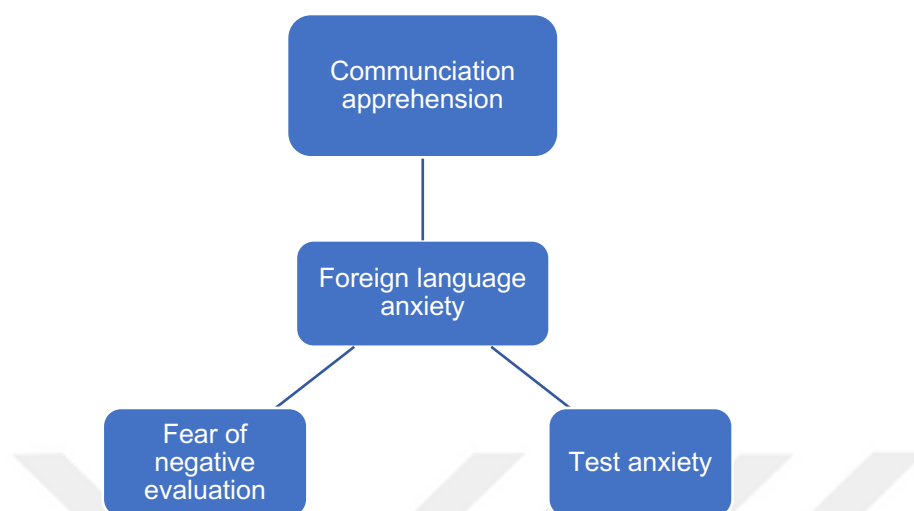
Young (1991, p.427) attempted to identify six general origins of language anxiety after a thorough review of the literature on the subject:

- Personal and interpersonal anxieties,
- Learner beliefs about language learning,
- Instructor beliefs about language teaching,
- Instructor-learner interactions,
- Classroom procedures,
- Language testing.

As can be seen from the list above, there are numerous possible causes of language anxiety. Some explanations relate to the pupil's emotional state, preferences, personality, or prior experiences. The instructor is a factor in other justifications. Other elements that are thought to contribute to language anxiety in language classes include how the language learning process is carried out in the classroom, the kinds of practices followed, and testing. Furthermore, the combination of the aforementioned causes may cause language anxiety to manifest itself.

In addition to his definition of foreign language anxiety, Horwitz et al. (1986) also suggested that language anxiety stems from three sources: (1) communication apprehension, (2) fear of negative evaluation by others, and (3) test anxiety. He also reported that even though they are useful conceptual cornerstones for defining foreign language anxiety, it is not merely a collection of these concerns applied to foreign language learning.

Components of Foreign Language Anxiety: Horwitz's Classification of FLA. As foreign language anxiety is concerned with performance evaluation in an academic and social environment, it's relevant to compare it to three other performance anxieties (Horwitz, Horwitz & Cope, 1986): Communication apprehension, test anxiety, and fear of negative evaluation.

Figure 3*Components of Foreign Language Anxiety*

Communication Apprehension. Communication apprehension is a kind of timidity marked by apprehension or fear of communicating with others (Horwitz, Horwitz & Cope, 1986). Thus, it can also be named performance anxiety or communication anxiety. Communication apprehension is crucial in language learning because communication plays an essential role in foreign language learning. People who struggle to talk in groups are more likely to struggle even more in a foreign language class, where they have less control over the communicative environment and their performance is continually assessed (Horwitz, Horwitz & Cope, 1986). It's likely that those people's incapacity to explain themselves or comprehend others via exchanging meaning might cause them more trouble. In addition to all of the usual problems regarding spoken communication, the foreign language class demands the student to communicate through a medium in which they have only limited proficiency (Horwitz, Horwitz & Cope, 1986). In other words, one of the origins of communication anxiety could be the necessity of generating language structures in a language that is not yet fully learned.

Test Anxiety. It has always been emphasized how important testing is. As a result, it's not unexpected that most students are anxious about testing because they are put in a

situation where their knowledge and abilities are examined in a limited amount of time. Test anxiety is a well-studied personality trait, in part because it gives a metric for the psychological importance of one specific type of stressful situation: those in which success is measured (Sarason, 1984). Students who are test-averse in a foreign language class are likely to have issues because tests and quizzes are prevalent in the language learning process, and even the brightest and best-prepared students can make mistakes (Horwitz, Horwitz & Cope, 1986). As a consequence, it's probable that test anxiety is primarily motivated by a fear of failing. Students who experience this kind of anxiety can also have difficulty learning and recalling content during tests, resulting in low test performance as test anxiety is a type of performance anxiety (Čiček, 2015).

Fear of Negative Evaluation. Since it appears in a variety of social circumstances in which a person may experience anxiety, the fear of negative evaluation is also associated with social anxiety (Čiček, 2015). Thus, it can occur in any social and evaluative situation (Horwitz, Horwitz & Cope, 1986). Both social anxiety and fear of negative evaluation seem related to an individual's self-esteem. People who struggle with social anxiety are not seen to have strong self-esteem. They are inclined to limit their social interaction because they are concerned about others' impressions of them and the potential of not being accepted socially.

Fear of negative evaluation is also related to test anxiety in that it is not restricted to test-taking scenarios; rather, it can appear in any evaluative scenario, for as in a foreign language lesson while students are speaking (Horwitz, Horwitz & Cope, 1986). In this sense, it can be conclusively stated that, while fear of negative evaluation can manifest itself in any situation, it is especially common in testing situations when the teacher is evaluating a written or spoken task because it focuses on the possibility that others will negatively evaluate them. Horwitz et al. (1986) also emphasized that anxiety over learning a new language often manifests itself in testing situations. Students may be concerned about being negatively judged by their teachers or peers as a result of the possibility of making

mistakes, and as a consequence, they may minimize or limit their efforts in order to avoid unfavorable feedback.

Effects of Anxiety in Language Learning Process

Many researchers are concerned that language anxiety may operate as an affective filter, preventing a learner from achieving a high level of competency in a foreign language (Aida, 1994; Cheng, 2002; Horwitz, 1986; Tuncer & Doğan, 2015). In FL learning, Aragão (2011) examined the relation between emotions and anxiety-provoking beliefs. He claimed that views about a student's self-concept influence feelings including shyness, embarrassment, self-esteem, and suppression. He also added that a student's worry about feeling embarrassed when speaking in class may stem from a fear that a classmate will criticize or scoff at his/her performance. In this sense, it's possible that unfavorable preconceptions about a student's own performance are linked to a negative self-perception that can result from a lack of confidence. Čiček (2015) also concluded that students' self-esteem will be much lower if they have a negative perception of their academic development, which will have an unfavorable impact on their learning. Cheng et al. (1999) further point out that low self-esteem appears to be a crucial component of the anxiety construct in this regard.

The above-mentioned possible negative effects of anxiety also accord with Čiček's (2015) study, which showed that anxiety could lead to poor performance when the language-learning circumstance is a source of stress and as the student's attention is drawn to the perceived risk rather than the language. In addition to the points mentioned above, one of the most prominent determinants of self-confidence or self-perceptions of competence is the absence of anxiety (Cheng, 2002). In other words, it is easier to foster self-confidence when learners are involved in activities that provide them a sense of accomplishment and enjoyment.

From a contrary perspective, some studies suggest that anxiety can be a motivating and contributing component in language classes. Students can utilize the language anxiety

instrument to perform better and gain more approval if they know how to manage their anxieties and trepidation while fulfilling their responsibilities. In other words, language anxiety, in this sense, encourages and motivates language learners emotionally to deal with challenging or unfamiliar tasks that might push them. Thus, anxiety serves as a motivator for language learners as a facilitator component in the learning process (Gardner & MacIntyre, 1993; Scovel, 1991; Young, 1991).

Ways to Decrease Anxiety

Anxiety in the classroom, whatever its source, is almost always an unwelcome circumstance (Tuncer & Doğan, 2015). That is to say, there is a range of sources that can be named as internal and external factors such as learners' ideas about themselves as individuals and language learning, teacher-learner relationships, exams, and in-class activities and these can create language anxiety. In this scenario, anxiety can have a variety of effects on students; while it might be counterproductive in certain cases, it can also stimulate them to improve. On the other hand, it is usually an unfavorable element since its possible negative impacts can interfere with language learning by distracting learners through various inner factors associated with anxiety, such as stress or fear. As a result, Huang (2012, as cited in Tuncer & Doğan, 2015) concluded that it has long been emphasized in the literature that anxiety should be decreased in order to facilitate learners' language learning process.

Arnold (2011) discussed that the relationship between learning and effects which refers to individual or personality factors (self-concept/self-esteem, anxiety, inhibition, and attitudes) has an impact on any classroom situation. She also addressed that this is particularly significant in language learning because our self-image is more vulnerable when we lack mastery of our vehicle for expression – language. Therefore, it will be simpler for language learners to deal with the constant risks associated with the language learning process if the components of the language learning process, such as in-class practices and testing concerns, are regarded as a pleasurable experience (Chneg, 2002). To wrap up, it

appears critical to create a learning atmosphere in which students may perform without experiencing embarrassment, anxiety, or fear. Arnold (2011) also suggested that:

- Language learning should operate in a low-anxiety environment.
- Classroom activities should include opportunities for students to succeed and therefore build confidence.
- The learner's cognitive, emotional, and physical components should all be considered.

As a basis of what Arnold indicated, language teachers should consider emotional elements and place a premium on dealing with any negative self-beliefs or attitudes that may hinder students' progression.

Language anxiety has also been linked to instructor ideas about language teaching (Čiček, 2015). Students' anxiety levels will rise if the teacher is unwilling to adjust or modify his or her teaching style which is based on his or her views of students with language learning difficulties (Piechurska-Kuciel, 2008 as cited in Čiček, 2015). In this respect, instructors are accountable for their students' language learning experiences, and as a result, they should invest as much time as possible in creating a stress-free and enjoyable learning environment. As a result, it seems crucial to recognize negative feelings, investigate their sources, and take actions to transform them could be functional to facilitate a positive classroom climate. Aida (1994) also confirmed the importance of instructors in the learning process. She reported that teachers play an essential role in reducing classroom tension and fostering a friendly, supportive environment that can help students overcome their fear of being embarrassed if they make mistakes in front of their peers. "If the teacher is too strict in error correction and does not adjust its intensity and frequency, the students may feel intimidated and experience negative emotions and anxiety, which is certainly not desirable in a language learning environment" (Čiček, 2015, p.13). Thus, it can be concluded that anxiety is also linked with language teachers. Students will value and learn more from teachers who can detect students who are anxious about learning a foreign language and take appropriate steps to assist them to overcome their fears (Aida, 1994).

Alrabai (2015, p.11) confirms the importance of language teachers by offering the following suggestions: Language teachers can help students overcome their fears of communication, negative evaluation, and language testing, appropriately address anxiety-provoking beliefs and misconceptions, assist students set clear and realistic goals for learning English, and boost their self-confidence.

In view of all that has been mentioned so far, one may suppose that with the help of an unthreatening and supporting learning setting, students may manage their anxiety more simply and efficiently. It also appears that a more precise diagnosis of learners' anxiety problems is a fundamental requirement for more effective interventions, as anxiety can obstruct effective learning. As Zheng (2008) points out that studying students' negative attitudes, as well as their causes and effects, can help educators devise effective interventions to help foreign language learners turn their negative views into positive ones.

Positive Psychology and Foreign Language Enjoyment

Although the involvement of positive emotion is generally acknowledged in the field, it still has a long way to go before it is accorded the recognition it deserves (Gardner, 1985; Krashen, 1982), because it appears to have stayed relatively in the shadows of the thriving research into negative emotions such as anxiety, which is mostly conducted in foreign languages (Dewaele et al., 2017). Positive Psychology, the empirical study of how individuals thrive and prosper, may be having an impact on the issue. Positive Psychology seeks to expand the general overview in psychology by focusing on the development of tools to improve positive emotions, enhance higher levels of engagement, and increase appreciation of meaning in life and its activities, rather than on the development of ways to reduce fear and learn to cope with negative experiences (MacIntyre & Mercer, 2014).

Many language educators recognize the importance of enhancing learners' language learning experiences by assisting them in developing and maintaining persistence and endurance, as well as positive emotions (MacIntyre & Mercer, 2014). The

aforementioned emotions and feelings may be necessary for the long-term endeavor of learning a foreign language because the language learning process is relatively long and includes being tested, which is an inextricable part of the language learning journey in order to see how effective the learning process is. As a result, while unpleasant feelings become more noticeable in language classes, positive and joyous feelings fade into the background. That's why, despite the fact that researchers have been studying the impacts and emotions regarding language learning since the 1970s, they have concentrated on negative ones like language anxiety (Dewaele et al. 2017).

On the other hand, it is accepted that learning a language takes place in a social environment. As learners internalize components of the context in which they live, the social context in which they learn is a crucial source of both positive and negative attitudes (MacIntyre & Mercer, 2014). The idea that negative emotions like fear, anxiety, boredom, or self-doubt can impede language learning and that figuring out how to get rid of them can help thoroughly clean out the longstanding effects of negative arousal is supported by the idea that not only getting rid of negative emotions but also nourishing positive emotions can make a stronger difference.

Botes et al. (2021b) also discussed that the association between foreign language enjoyment and anxiety is most likely cyclical, as less anxiety can lead to more enjoyment in the FL classroom, which can lead to reduced anxiety. As MacIntyre and Mercer (2014) stated that negative emotion tends to constrain a person's circle of attention and predisposes specific action tendencies, whereas positive emotion fosters dispositions toward self-exploration, extending a person's range of concentration and developing resources for future action. Furthermore, positive emotions, according to MacIntyre and Gregersen (2012), have consequences that go beyond pleasant feelings: it improves learners' ability to perceive items in the classroom environment and strengthens their awareness of language input (as cited in Dewaele et al., 2017). It is therefore also appropriate to take enjoyment into consideration while assessing attitudes towards

assessment and evaluation instruments that are used in testing where the language input provided to the students is measured.

With the growing interest in the importance of emotions in language learning, a more holistic approach to learners' emotions has led to the incorporation of positive emotions in research designs. The idea that researchers should not just focus on what is wrong, but also on what is good and generates pleasure emerges as a result of the complex dynamics of emotions (Dewaele et al., 2019). The foreign language enjoyment scale to evaluate positive feelings in language learning, with items such as "I enjoy..." or "There is a positive..." as well as word choices in statements like "enjoy, like, fun, interesting, encouraging, supportive, and friendly" was developed by Dewaele and MacIntyre (2014). They define factors in FL learning that incorporate positive emotions into three categories (Botes et al., 2021a):

- Teacher appreciation, which emphasizes the teacher's positive attitude or characteristics;
- Personal enjoyment, which is linked to the student's own personal choice and positive feelings about his or her achievements,
- Social enjoyment, which includes positive feelings about the rest of the classroom, such as peers and instructors as well as the individuals' sense of community.

To conclude, the concept of foreign language enjoyment is crucial in shifting the focus of language learning away from negative feelings and toward all aspects of the process, including materials, people (peers, teachers, and tests), and other tools that promote positive attitudes (Dewaele et al., 2017). Furthermore, emphasizing positive emotions can broaden students' social, personal, and intellectual resources, as well as motivate them through a difficult language learning process and increase their awareness of what they have accomplished.

Recent Studies

Studies on the Attitudes of Students/Teachers on Language Assessment

Gökhan (2004) conducted a study to find out how teachers at a public university's Basic English Division felt about the current assessment system in place. The findings revealed that both teachers and students considered the assessment method satisfying in terms of its own merits as well as its shortcomings and strengths. Another finding of the study is that teachers want to be involved in the assessment process by using alternative assessments, and students want to be involved by offering their comments on tests and test activities.

Struyven et al. (2005) conducted research with the goal of examining evaluation and assessment from the perspective of students. According to the findings, students' attitudes toward evaluation have a substantial impact on how they learn and study. Students' study methods, on the other hand, have an impact on how they view evaluation and assessment. According to the findings, students have strong feelings about various assessment and evaluation formats and they prefer multiple-choice examinations to essay-style questions in this regard. They, on the other hand, criticize the 'fairness' of these well-known evaluation techniques when compared to more novel assessment approaches.

Paker and Höl (2012) conducted a study to investigate students' and instructors' attitudes and perceptions of the speaking test at a School of Foreign Languages. The statistics revealed that the majority of the students had never taken a speaking test before, hence they were more nervous during the test. When compared to the testing of other language abilities, the speaking test was viewed as the most challenging by the students. Students stated that they were unable to fully express themselves during the test. The instructors, on the other hand, stated that the speaking test was the most difficult to administer and evaluate; however, the scale and rubrics were appropriate in evaluating the students' oral performance.

Yüce (2015) conducted a study to evaluate pre-service English language teachers' views of assessment and to determine which assessment approaches they envision utilizing when they begin teaching English. The findings showed that while pre-service English language teachers see evaluation as a tool to improve the quality of instruction and student learning, measure the quality of schools, and certify students' progress, most pre-service English language teachers also see assessment as irrelevant. Furthermore, the descriptive data suggested that pre-service English language teachers are more likely than their language teachers to consider alternate assessment practices.

Deretarası (2020) conducted research at state universities in Turkey to determine the knowledge level of in-service EFL instructors in general and skill-based language assessment, as well as the impact of gender, years of work experience, and experience in an assessment unit on English instructors' assessment knowledge levels. The findings revealed that the participants' understanding of general and skill-based language testing and evaluation is low. According to the findings, having work experience as a member of an assessment unit has a statistically significant impact on the assessment expertise of EFL instructors. Gender differences have no statistically significant impact on the teachers' evaluation knowledge levels. There was also no relationship discovered between years of job experience and assessment knowledge level.

Studies on the level of FLA/FLE in Language Learning

Wörde (2003) examined what elements students thought contributed to anxiety and how to lessen it. According to the interviews and replies to the FLCAS questions, 73 percent (11) of the 15 participants in the study were anxious learners, with 34 percent (5) of them being extremely anxious. The interviews suggested that the language class was exceedingly unpleasant. Speaking activities, difficulty to comprehend, poor classroom experiences, fear of negative evaluation, native speakers, methodology, pedagogical practices, and the teachers themselves were cited by the participants. Speaking and listening activities were found to be the most anxiety-inducing variables.

Aydin (2008) conducted a study into the sources and levels of fear of negative evaluation, as well as language anxiety, among Turkish students who are learning English as a second language, and the relationship between the two. A total of 112 foreign language learners were given a foreign language anxiety scale and a scale for fear of negative evaluation. Descriptive and correlational analysis has been performed using the data collected. The findings of the study revealed that EFL students experience linguistic anxiety and fear of negative evaluation. Furthermore, language anxiety was discovered to be a prominent source of the fear of negative evaluation.

Trifoni and Shahini (2011) conducted research on test anxiety and its consequences on learning, as well as the origins and effects on students. The findings revealed that a large number of students were impacted by test anxiety, at least to some extent. Some of the variables that contribute to test anxiety include a lack of or insufficient test preparation, fear of negative evaluation, previous test failures, time constraints and pressure, the number of items on the test, and the difficulty of the course subject. Test anxiety causes both physical and psychological difficulties, according to the findings. It has a detrimental impact on motivation, focus, and achievement, as well as increasing exam mistakes, making it difficult to recall previously taught content, and preventing efficient study.

Demirdağ (2012) conducted research to see if there was a relationship between foreign language anxiety levels and language performance among students at a University English Preparatory School with an A1 English proficiency level. The study's findings revealed that the participants in their English classes had comparatively low levels of anxiousness. There is a substantial significant relationship between gender and language anxiety, but not between gender and language performance, according to the findings. It is also revealed that whereas age and language performance are highly associated, age and language anxiety are not, and there is a significant negative relationship between foreign language anxiety and language performance.

Čiček (2014) studied three groups of young learners from two different primary schools to see if foreign language learning anxiety exists from the start of the language learning process. Students reported having a low or very low level of anxiety, confirming the prevalent belief that young learners have very little anxiety during their language learning process. There were substantial disparities in anxiety levels between male and female students: Female students were more anxious than male students. Another important conclusion was that language anxiety and motivation were negatively correlated.

Tuncer and Doğan (2015) conducted research to determine to what extent Turkish students' English classroom anxiety influences their academic performance in the English language. Foreign language anxiety at the start of prep-class education, according to the research, was ineffective as a predictor of academic achievement. On the other hand, it was discovered that students' anxiety at the end of their education predicted academic performance with their interest in the language class dimension and accounted for academic failure with their speaking anxiety in the language class dimension. They discovered that the students' anxiety levels fluctuated and increased throughout their English prep school and that this fluctuation predicted their academic achievement.

By uncovering the reasons for FLA in preparation school students in English courses and identifying solutions to lessen or overcome FLA in preparatory school students, Ziyen Atlı (2017) addressed the subject of foreign language anxiety (FLA) from the perspectives of both foreign language learners and teachers. The study's findings revealed that there is a statistical difference in communication apprehension between the courses. Similarly, significant differences in the anxiety of negative evaluation were discovered throughout the classes. In terms of exam anxiety and classroom procedure, however, there were no significant differences across the classes. The study also revealed some of the origins of foreign language anxiety, as well as some strategies for reducing or eliminating it.

Mierzwa (2018) examined the impact of foreign language enjoyment on learning English as a foreign language, as well as the gender implications of this relationship. The

findings of this study revealed that there are no statistically significant differences in FLE between males and females, but there are differences in which sources of FLE each gender considers to be the most important. It has been proven that FLE tends to grow with a student's competency level, and that a high degree of FLE leads to higher academic accomplishment.

Dewaele et al. (2019) conducted a study of 592 Turkish as a foreign language (FL) learners in Kazakhstan, focusing on foreign language enjoyment (FLE) and foreign language classroom anxiety (FLCA). FLE and FLCA mean levels were found to be similar, with a weak positive correlation between the two. Male individuals reported higher FLCA than female participants. Multiple regression analyses demonstrated that learners' attitudes toward Turkish and teacher-related variables predicted FLE and FLCA more strongly than learner-internal variables, supporting earlier research outside Kazakhstan. A total of 25% of the variance in FLE was explained by attitudes toward the FL, teacher friendliness, strictness and frequency of usage of the FL, attitude toward the teacher, participant's age, and FL exam outcome. The data show that FLE and FLCA variation among Kazakh Turkish learners is similar to that observed in other contexts.

Li (2020) looked into the complicated relationships between Trait Emotional Intelligence (TEI), Foreign Language Enjoyment (FLE), and English-as-a-foreign-language (EFL) learning achievement in 1307 Chinese high school students. He discovered that the majority of Chinese high school EFL students had moderate to high TEI and low to moderate FLE. Students' TEI, FLE, self-perceived English achievement, and actual English achievement all had low to medium correlations. Finally, FLE-mediated TEI indirectly influences perceived achievement and actual achievement.

Kayhan (2021) employed mixed methods research to investigate Foreign Language Anxiety (FLA) among young learners by querying the level of FLA among young learners and explaining the sources of FLA among young learners. According to the findings, the young students in the class exhibited a moderate level of FLA. Age, gender, academic

accomplishment, parental education level, and parental expectation all had an effect on young learners' FLA, according to the study. In addition, FLA among young learners is caused by unexpected questions from teachers, inability to answer the questions, poor grades, teachers, challenging subjects, writing long articles, and some classroom activities.

Özer and Altay (2021) investigated the role of positive and negative emotions in language learning in terms of enjoyment and anxiety, two emotions that have been studied extensively in the field of language education. Participants indicated a higher level of enjoyment and moderate anxiety in language learning, according to the findings. Gender differences did not appear to alter participants' anxiety or enjoyment levels. In terms of anxiety and enjoyment, there was a significant difference between the three achievement groups. Language anxiety was found to be the best predictor of language enjoyment after further investigation.

Chapter 3

Methodology

This chapter is aimed at presenting information regarding the research design, participants and instruments used to collect data along with data collection techniques and data analysis.

Types of Research

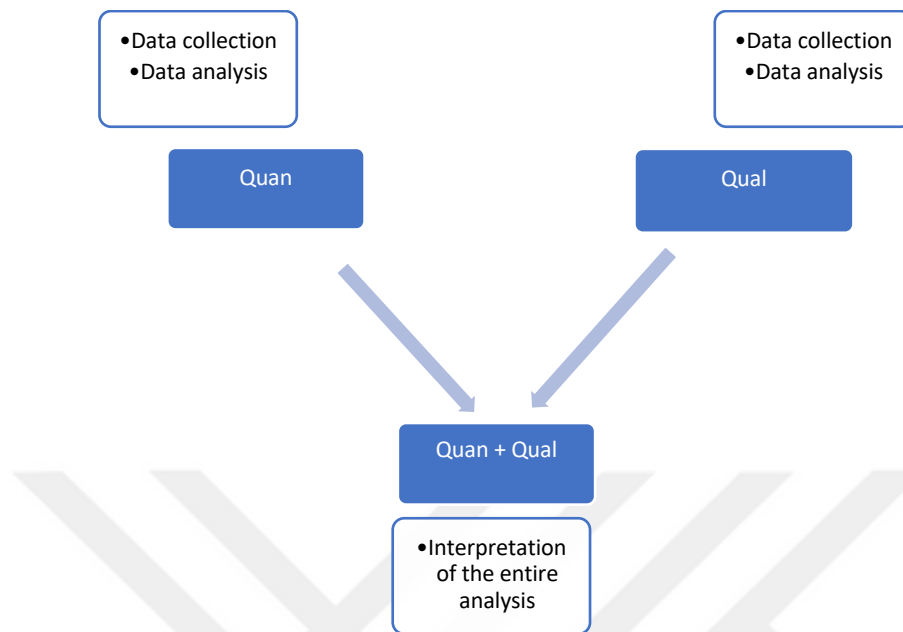
A survey method that includes collecting and analyzing both quantitative and qualitative data in order to uncover the attitudes of A1-level preparatory school students toward assessment and evaluation instruments by determining their anxiety and enjoyment levels in the context of positive and negative attitudes is used in this study.

In order to gain a thorough understanding of assessment and evaluation instruments from the viewpoints of students and to shed light on certain characteristics of a particular instance, this study also examines the testing practices at a particular university. Since it seeks to investigate a situation in its context to produce a multifaceted understanding of a complicated issue, this study can therefore be viewed as a case study. According to Crowe et al. (2011), the case study approach, a well-established research approach that is widely utilized in a wide range of disciplines, particularly in the social sciences, allows for complete, in-depth studies of complex issues in the context of real-world events.

The current study, which is also a case study, makes use of the survey method by giving priority to gathering and analyzing both quantitative and qualitative data at the same time in order to cross-check and validate the findings of each way of data collection and analysis. The fact that it might help the researchers provide a greater insight than they would have from either qualitative or quantitative research on its own is another reason why this research strategy was chosen for the current study. The research design for the current study is displayed in Figure 4 below.

Figure 4

The Research Design of the Study (Adapted From Creswell, 2009)



Setting and Participants

The study was conducted with A1-level students studying at preparatory school at a public university in Ankara. Generally in the program of preparatory departments of the universities, alternative assessment strategies such as portfolios are used, as well as more standard instruments such as quizzes and exams. Students must obtain a legitimate score before taking the proficiency exam at the end of the year, which will determine whether they pass or fail. It is common that the testers are in charge of creating the tests, while the instructors are in charge of administering and scoring them.

In the department of Basic English, where this study takes place, four primary types of assessment tools are used. Assessment instruments can be classified into two types based on the assessment tools used in this department: Hard ones that refer to a more traditional way of assessing that focuses on the product, and soft ones that are more flexible and emphasize the process in which the student's learning process is taken into account. The detailed description of each instrument can also enhance our understanding of

assessment and evaluation instruments that are the core component of this study and offer a framework for investigating students' attitudes toward them:

- *The speaking assessment/test* includes task/tasks requiring using the key language structures, other useful phrases, and vocabulary items in the coursebook and the program. It is administered as an in-class speaking activity with a grading scale for evaluating each student's performance. The speaking assessment might be dealt with as a hard assessment instrument as it focuses on the product. However, it is possible that it encompasses only a few occurrences of soft assessment because it is qualitative in character and aims to make the learning process more democratic by providing an alternate method of evaluating students' work (Frodden Armstrong et al., 2004).
- *Exams* such as quizzes are commonly used to examine a variety of components of the language, such as grammar (also known as language), reading, and vocabulary. Students are notified ahead of time of which components of the language will be evaluated in the quiz. Mid-term exams, unlike quizzes, include language, receptive, and productive abilities, such as writing, to assess students' progress through the coursebook and program. Quizzes and exams are examples of hard assessment instruments that are given at the beginning, middle, and end of the semester as part of a summative evaluation.
- *The writing portfolio* is comprised of various writing tasks that correspond to the content and language studied during the term. Students are expected to produce two drafts, the first during the program's allowed class hours and the second/final draft after receiving comments from the teacher outside of class. The teachers grade each task using a rubric provided by the department.
- *The online components* include the online practice of the assigned exercises from the allotted units that students must complete. Students can complete the online exercises multiple times and receive feedback. Students' writing portfolios and online component scores are reviewed and incorporated into their semester average. As a result, soft assessment instruments such as writing portfolios and tasks in online components are considered part of continuous assessment.

There is evidence that students favor whichever system of evaluation they are exposed to, regardless of its benefits or flaws (Blaikies, Schönau, & Steers, 2004). As a result, “students may not really be in a position to evaluate assessment methods that are not in their experience” (p.64), which means they simply normalize and adapt to the institution's preferences. It is important, however, to look at their attitudes toward assessment and evaluation procedures. Because assessment of learning is such an important element of the teaching-learning process in any discipline, it should be viewed as both an evaluation and a learning activity (Moore, 2010).

In the quantitative part of the study, the target population of the study was A1-level students who were studying in the preparatory class in the 2021-2022 academic year fall semester. “The aim of all quantitative sampling approaches is to draw a representative sample from the population so that the results of studying the sample can then be generalized back to the population” (Marshall, 1996, p. 522). Based on Marshall's statement, A1-level students whose population outnumbers the population of other levels were preferred. The participants were chosen according to random sampling and the number of participants who answered the questionnaire was 218. In the first part of the questionnaire, the participants were asked to choose their gender and academic major. The number of female participants was 104 (47,7%) and the number of male participants was 114 (52,3%). The academic majors that students will study in were as follows: Social Sciences: 121, Natural and Applied Sciences: 97 (See Table 1)

Table 1

Descriptive Statistics of the Participant Students in the Quantitative Part

		Frequency	Percent
Gender	Female	104	47,7
	Male	114	52,3
Academic Major	Social Sciences	121	55,5
	Natural and Applied Sciences	97	44,5
Total		218	100,0

In the qualitative part of the study, convenience sampling that allows researchers to select subjects who are easily available and convenient to them was preferred (Baltacı, 2018; Creswell, 2009). This method of sampling does not necessitate a random selection of participants based on any list of requirements, instead, the researcher can select participants at random who are willing to be contacted and participate in the study.

The students who participated in semi-structured interviews were selected according to convenience sampling which is a purposive sampling method from A1-level university preparatory class students studying at a public university in Ankara. In the 2021-2022 fall semester, university preparatory class students studying at a public university in Ankara were asked to attend the research by taking their consent. In total, five female (50,0%) and five male (50,0%) participant students were involved in this study. The academic majors that students will study in were as follows: Social Sciences: 8, Natural and Applied Sciences: 2. The following table includes detailed information about the students who attended this study.

Table 2

Descriptive Statistics of the Participant Students in the Qualitative Part

Participant Students	Gender	Academic Major
1	Male	Social Sciences
2	Female	Social Sciences
3	Female	Social Sciences
4	Male	Natural and Applied Sciences
5	Female	Natural and Applied Sciences
6	Male	Social Sciences
7	Male	Social Sciences
8	Male	Social Sciences
9	Female	Social Sciences
10	Female	Social Sciences

Data Collection

The purpose of this study was to investigate the attitudes of A1-level university preparatory school students toward the various forms of assessment and evaluation instruments employed in the department. The data for this study were collected using a questionnaire created by the researchers. The instrument consisted of two sections and a total of 32 statements within the Likert format. The first section, Part I, consisted of 2 questions that asks the gender and academic major of the participants. The second section, Part II, consisted of 32 statements aimed at determining students' attitudes by finding out whether the students feel anxious about or enjoy various types of assessment and evaluation instruments (See Appendix F).

Instruments

In the present research, one of the instruments which aim to discover the attitudes of English language learners studying at a university preparatory school towards different assessment and evaluation instruments is the questionnaire that was created by the researchers. The reason why a questionnaire is used to investigate attitudes in the current study is that individual attitudes are particularly important to track since they are thought to be precursors to conscious actions and questionnaire responses are the most popular method of measuring attitudes (Swamy, 2007). A semi-structured interview form that was formed by the researchers was also used to get the more detailed information regarding the research question.

Instrument 1

Instruments meant to measure language anxiety and enjoyment are dominated by items covering four skills of language achievement or performance, primarily in-class activities (Cheng et al.,1999). However, the current study focuses on assessment and evaluation instruments. Thus, a comprehensive literature review was used to create the item pool for the questionnaire, and the items of "The Foreign Language Classroom Anxiety

Scale" (FLCAS) developed by Horwitz et al. (1986) and "The Foreign Language Enjoyment Scale" (FLES) developed by Dewaele and MacIntyre (2014) were thoroughly examined. Following the creation of the questionnaire's items, the content validity of the items was confirmed by consulting two experts in the fields of language teaching and educational sciences, and the items were rearranged and five items were removed and the expressions in several questionnaire items were simplified to make it easier for participants to grasp the items and avoid confusion based on expert recommendations in the questionnaire used by the researcher in this study.

Literature also suggests that pretesting that is done by means of a personal interview can be a good way to detect ambiguous statements and ambiguous word meanings (Altunışık, 2008; Reynolds, et. al, 1993). Thus, it is important to determine the respondents' reactions to the questionnaire. According to Reynolds et al.. (1993), the technique for determining the respondents' responses to the questionnaire can take one of two forms: the respondent can think aloud while filling out the questionnaire, or the respondent can talk about the questionnaire after it is completed. In this study, the latter was preferred and the opinions of some of the participants that are A1-level students were collected after completing the questionnaire. When participants were kindly asked to fill out the questionnaire, they could see the items in English along with Turkish equivalence since the participants are A1-level students. The participants who took part in the piloting stated that they could easily understand the statements and there were no confusing or loaded statements. In addition, according to the participants in the piloting, the options ranging from "strongly disagree" to "strongly agree" looked enough for expressing their thoughts.

During a pandemic, where students attend classes from distant cities, an online questionnaire seemed to be the ideal data collection method. The study was based on a Google platform application that includes a structured five-point Likert questionnaire that can be completed online. The questionnaire is titled: *"University English Preparatory School Students' Attitudes Towards Assessment and Evaluation Instruments"*.

Table 3*Questionnaire Items*

Item numbers	
Items regarding anxiety	Assessment & Evaluation Instrument
1-6	Speaking Assessment
7-11	Exams (Quizzes & Mid-term exams)
12-15	Writing Portfolio
16-18	Online Components
Items regarding enjoyment	
19-24	Speaking Assessment
25-27	Exams (Quizzes & Mid-term exams)
28-29	Writing Portfolio
30-32	Online Components

The questionnaire which aims to discover the attitudes of English language learners studying at a university preparatory school towards different assessment and evaluation instruments involved 32 items with two separate sections including 17 statements related to anxiety and 15 statements related to the enjoyment of different assessment and evaluation instruments. The participants were presented with five options – *Strongly Agree*, *Agree*, *Neither Agree Nor Disagree*, *Disagree*, *Strongly Disagree* – next to each item and they were expected to choose the option that best explains their views.

According to Taber (2017, p.1278), “Alpha was found to be used as an indicator of reliability”. Cronbach's alpha is also commonly employed in the design of scales and questionnaires. Taber (2017,p. 1278) stated that alpha values were described as “high (0.73–0.95), good (0.71–0.91), and relatively high (0.70–0.77)” by different researchers. Thus, the reliability of the questionnaire was determined in the current study using Cronbach's alpha, and it was determined to be reliable ($R=.73$). (**Table 4**).

Table 4*Reliability Analysis of the Questionnaire*

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.736	.721	32

Instrument 2

As it's critical to understand how assessment tools affect students' anxiety and enjoyment levels, as well as to identify the sources of students' perceptual differences about assessment tools, an interview form was created to gather more information and double-check the results obtained from the questionnaire. Thus, a semi-structured interview form was used to determine how students feel about assessment and evaluation tools, as well as the origins of students' attitudes about assessment instruments, which had 4 main questions and 3 sub-questions for each (See Appendix G). The interview questions were created objectively based on thorough literature research and expert opinion on a variety of issues, including the necessity and relevance of individual questions, as well as loaded or leading questions. The procedure of consulting the participants' ideas was also used to see whether there were any unclear expressions so that the interview form could be changed accordingly. As a result, some minor changes have been made in the questions in terms of language and the interview questions were found to be applicable and convenient to apply.

The semi-structured interview form was administered to 10 A1-level students who were studying in the preparatory classes and took part in the quantitative part of the study. Ten (10) separate semi- interviews were conducted in Turkish which is the native language of the participants as they are A1-level students. "A semi-structured interview is a meeting in which the interviewer does not strictly follow a formalized list of questions. They will ask more open-ended questions, allowing for a discussion with the interviewee rather than a straightforward question and answer format." (Doyle, 2019, p.1). In these interviews,

participants were informed about the study and an online consent form was signed by the participants within the scope of the study, and open-ended questions about the assessment and evaluation instruments were asked by the researcher. Webex, a tool that allows users to participate in video conferencing, was used with the permission of the participants. Interview notes were also kept by the researcher. Tablo 5 shows the details of the data collection instruments used in the current study.



Table 5*Data Collection Instruments*

Research questions	Data collection method	Quality of the participants	Number of the participants	Instruments	Data analysis
1.What are the attitudes of A1-level students toward assessment and evaluation instruments?	Quantitative	A1-level prep class students	218	Questionnaire	Descriptive statistics
2. Which of the assessment and evaluation instruments causes the highest and the lowest anxiety among A1-level students?	Quantitative	A1-level prep class students	218	Questionnaire	Descriptive statistics
3. Which of the assessment and evaluation instruments causes the highest and the lowest enjoyment among A1-level students?	Quantitative	A1-level prep class students	218	Questionnaire	Descriptive statistics
4. Are there any significant differences in the attitudes of A1-level students regarding the extent of anxiety and enjoyment in terms of gender?	Quantitative	A1-level prep class students	218	Questionnaire	T-test
5. Are there any significant differences in the attitudes of A1-level students regarding the extent of anxiety and enjoyment in terms of academic major?	Quantitative	A1-level prep class students	218	Questionnaire	T-test
6. What factors emerge from speaking exams that affect participants' attitudes in the context of anxiety and enjoyment?	Qualitative	A1-level prep class students	10	Semi-structured interview form	Content analysis
7. What factors emerge from quizzes and mid-term exams that affect participants' attitudes in the context of anxiety and enjoyment?	Qualitative	A1-level prep class students	10	Semi-structured interview form	Content analysis
8. What factors emerge from the writing portfolio that affects participants' attitudes in the context of anxiety and enjoyment?	Qualitative	A1-level prep class students	10	Semi-structured interview form	Content analysis
9. What factors emerge from online components that affect participants' attitudes in the context of anxiety and enjoyment?	Qualitative	A1-level prep class students	10	Semi-structured interview form	Content analysis

Data Analysis

During the data analysis process of the quantitative research section, the data collected via the questionnaire were transferred from Google document to the SPSS 28. The reliability of the scale was measured with reliability analysis and found as .73. To begin, the participants' demographic information was presented in a table (see Table 1) that gives the total number of respondents together with frequencies and percentages indicating respondents' gender and academic major. Then, the data collected from the questionnaire were analyzed using descriptive statistics in the first phase to be able to describe the attitudes of the students towards different assessment and evaluation instruments. Moreover, to analyze the impact of gender and academic major, a t-test was conducted.

As equal variance and normality are assumptions in parametric mean comparison tests like the t-test and ANOVA, Orcan (2020) stated that skewness and kurtosis values should be verified. He also added that although there is no universal agreement on the numbers that indicate normality, values between -2 and +2 are commonly used to demonstrate normal distribution. The test of normality was run, and the results for skewness and kurtosis were found to be 0,14 and 0,08 (See Table 6), indicating that the values indicated a normal distribution based on the specified criteria.

Table 6

Test of Normality Results of the Questionnaire

Descriptives			
Total		Statistic	Std. Error
	Skewness	.140	.165
	Kurtosis	.085	.328

In terms of normality, the sample size is also crucial as normality of data cannot be quarantined with a small sample size (Orcan, 2020). The following table shows that a sample size of 217 students with a 95 percent confidence interval and a 5% margin of error

is required for the study (Krejcie & Morgan, 1970). According to data obtained from the department, 503 A1-level students are enrolled in English prep classes at the university where the study was conducted. In this regard, the number of students (218) who took part in the quantitative part of the current study is sufficient for the study.

Table 7

Required Sample Size for the Research

Population size	Confidence level = %95			Confidence level = %99		
	Margin of error			Margin of error		
	5%	2,5%	1%	5%	2,5%	1%
100	80	94	99	87	96	99
500	217	377	475	285	421	485
1.000	278	606	906	399	727	943
10.000	370	1332	4899	622	2098	6239
100.000	383	1513	8762	659	2585	14227
500.000	384	1532	9423	663	2640	16055
1.000.000	384	1534	9512	663	2647	16317
10.000.000	384	1536	9594	663	2653	16560

The descriptive analysis technique was utilized to summarize and evaluate the qualitative data (semi-structured student interviews) that were used to provide a broader, deeper, and an inside look into the assessment practices. Qualitative inquiry is based on a different philosophical assumption, which leads to diverse data analyses and interpretations by the researcher (Creswell, 2009). As a result, data analysis includes preparing data for analysis, scrutinizing the participants' words and sentences in order to have a better grasp of the data, and interpreting the data based on the themes or perspectives gleaned from the interviews. Thus, the qualitative data collected were analyzed through content analysis. Several components unique to the qualitative inquiry are included in the data analysis plan

(Creswell, 2009; Miles & Huberman, 1994). The following are the steps that were taken for data analysis:

- In order to organize and prepare data for analysis, raw data was transcribed.
- The researcher examined all of the data to get a general sense of it and to think about its overall meaning.
- According to the interview questions, the data was divided into segments, categories, and groups.
- To designate the themes with a keyword, the most descriptive phrasing for the topics was picked.
- Words and phrases that are recurring and significant in representing the themes are grouped and labeled. These labels are usually words and short phrases.
- Frequencies (f) were presented regarding the codes.
- Existing data was ready to be recoded if necessary, based on the expert's assessment.
- A hierarchical frame was formed to demonstrate how they relate to one another.
- The findings of the study were also presented as descriptive text with quotations from participants.
- The participants were indicated as P1, P2 and P3 under the relevant quotation.

Chapter 4

Findings, Comments and Discussion

The purpose of this chapter is to present the findings of the data analysis in relation to the research questions and to report the data acquired from the current study. The quantitative and qualitative components of the data analysis were carried out using various data analysis methods. The quantitative and qualitative data findings were presented separately.

Quantitative Data Analysis

Findings From the Questionnaire

The results of the questionnaire are presented in this section. The researcher used a questionnaire to analyze the attitudes of A1-level students regarding various assessment and evaluation instruments. The notions of anxiety and enjoyment—that is, what students are anxious about and what they enjoy—are utilized to determine whether students have positive or negative attitudes about four different assessment and evaluation instruments employed in the department. The SPSS 28.0 package program was used to analyze the data collected.

The descriptive analysis of the scale's items was presented. The minimum, maximum, standard deviations, and means of the questionnaire's items were evaluated, and Cronbach Alpha, the most generally used approach, was used to assess reliability. It was determined whether the data met the assumption of normalcy using the Skewness-Kurtosis coefficients, and it was discovered that the Skewness-Kurtosis coefficient for each item was in the range of 2. As a result, parametric techniques were applied in the data analysis. As a consequence, the difference between the average scores obtained from the scale according to gender and academic major was analyzed using an independent sample t-test. The level of statistical significance was set at 0.05 in the interpretation of the data as

the criterion for distinguishing significant from non-significant data has commonly been set at 0.05 (5 percent) (Leo & Sardanelli, 2020).

What are the attitudes of A1-level students towards assessment and evaluation instruments? The first research question investigates university A1-level students' attitudes about assessment and evaluation instruments. Table 4 shows the mean of the participants' total ratings from the anxiety and enjoyment items on the assessment and evaluation instruments.

Table 8

Analysis of the Average Mean Scores for the Anxiety and Enjoyment Items in the Questionnaire

	Descriptive Statistics				
	N	Minimum	Maximum	Mean	Std. Deviation
Anxiety items TOTAL	218	1.61	4.56	3.0780	.60018
Enjoyment items TOTAL	218	2.21	4.57	3.5056	.45568

Table 8 demonstrates that the mean of the anxiety items on the scale is 3,07, with 1 being the lowest and 4 being the highest. The average number of enjoyment items is 3,50. This section has a minimum score of 2 and a maximum score of 4. Since the means are similar, the findings suggest that students have both moderately positive and negative attitudes toward testing instruments rather than having a predominantly one-sided attitude. Nevertheless, despite the fact that the means are close, the mean of the enjoyment part is only modestly higher than the mean of the anxiety section. Thus, the participants' attitudes regarding assessment and evaluation instruments are probably slightly more positive.

In their study on FLE and FLA among Turkish EFL students, Özer and Altay (2021) discovered that participants experience a higher level of enjoyment (M=3,87) and moderate anxiety (M=3,19) in language learning. When the results of this study are compared to the findings of Özer and Altay's study, it is clear that participants in this study did not experience a high level of enjoyment even though participants in both pieces of research experienced

a similar level of anxiety. One possible explanation is that the current study emphasizes language testing, which is often regarded as a negative experience by students (Sarason, 1984), whereas the other study is focused on language learning in general which includes assessment and testing as components.

The mean scores of the items in the anxiety and enjoyment sections of the questionnaire are also shown in **Table 9** below.

Table 9

Descriptive Statistics of the Anxiety and Enjoyment Items in the Questionnaire

Questionnaire Items					
Items regarding Anxiety	N	Min	Max	Mean	S.D.
Speaking assessment					
1. Speaking exams make me nervous.	218	1,00	5,00	3,65	1,02
2. I feel stressed over not making proper sentences and explaining my ideas in speaking exams.	218	1,00	5,00	3,80	1,05
3. I feel nervous in speaking exams as I feel anxious over making a mistake in front of the instructor.	218	1,00	5,00	3,21	1,20
4. I feel stressed in speaking exams if the instructor isn't relaxed.	218	1,00	5,00	3,82	0,98
5. I feel nervous in speaking exams when we perform the task in pairs.	218	1,00	5,00	2,75	1,15
6. I feel stressed in speaking exams when I perform the task individually.	218	1,00	5,00	2,84	1,13
Exams					
7. Quizzes make me nervous.	218	1,00	5,00	2,77	1,17
8. Midterm exams make me nervous.	218	1,00	5,00	3,79	1,09
9. I feel anxious in quizzes and midterm exams as I experience anxiety-provoking self-talk. (e.g. I can't do it. I will forget everything etc.)	218	1,00	5,00	2,87	1,28
10. I feel stressed in quizzes and midterm exams because of the time limit.	218	1,00	5,00	3,36	1,24
11. I feel stressed in quizzes and mid-term exams if the instructor isn't relaxed.	218	1,00	5,00	3,49	1,12
Writing Portfolio					

12. Completing the tasks in the writing portfolio makes me nervous.	218	1,00	5,00	2,42	0,96
13. I feel nervous while completing the tasks in the writing portfolio as I have some linguistic problems (inadequate mastery of vocabulary and sentence structures, grammatical errors etc.)	218	1,00	5,00	2,90	1,11
14. I feel nervous about the negative comments and evaluation of the instructor about my work in the writing portfolio.	218	1,00	5,00	2,43	1,15
15. I feel stressed when I have to write under time constraints.	218	1,00	5,00	3,70	1,02
Online Components					
16. Doing the exercises in the online component (online platform of the coursebook) makes me nervous.	218	1,00	5,00	1,98	0,93
17. I feel anxious about doing the exercises in the online component as it is artificial intelligence that evaluates my work.	218	1,00	5,00	2,66	1,26
18. I feel stressed about completing the tasks in the online component as there is a deadline.	218	1,00	5,00	2,97	1,25
Items regarding Enjoyment					
Speaking Assessment					
19. I enjoy speaking exams.	218	1,00	5,00	2,88	1,13
20. I take pleasure in the speaking exam when the topic of the task is relevant to our interests.	218	1,00	5,00	4,01	0,90
21. I enjoy the speaking exam when we perform the task in pairs.	218	1,00	5,00	3,44	1,02
22. I enjoy the speaking exam when I perform the task individually.	218	1,00	5,00	3,11	0,99
23. I take pleasure in the speaking exam when I'm praised for my good performance.	218	1,00	5,00	4,19	0,88
24. I enjoy the speaking exam if the instructor's attitude is positive.	218	2,00	5,00	4,39	0,70
Exams					
25. I enjoy quizzes.	218	1,00	5,00	3,08	0,91
26. I enjoy mid-term exams.	218	1,00	5,00	2,39	0,88
27. I take pleasure in taking the quizzes and mid-term exams since they help me to identify my gaps.	218	1,00	5,00	3,45	0,98
Writing Portfolio					
28. I enjoy completing the tasks in writing portfolio.	218	1,00	5,00	3,55	0,87

29. I enjoy completing the tasks in writing portfolio as it helps me to improve my writing skills gradually.	218	1,00	5,00	3,80	0,85
Online Components					
30. I enjoy doing the exercises in the online component.	218	1,00	5,00	3,32	1,02
31. I am pleased by completing the tasks in the online component as I can do extra exercises.	218	1,00	5,00	3,62	0,92
32. I enjoy doing the exercises in online component as I can do them at my own pace.	218	1,00	5,00	3,84	0,92

When the mean values of the items that make up the anxiety section were examined, it was discovered that while an item from the speaking assessment section has the highest average on the whole scale, an item from the online components section has the lowest average mean. It was shown that while "I4- *I feel stressed in speaking exams if the instructor isn't relaxed.*" (M=3,82) has the highest average score, "I16. *Doing the exercises in the online component (online platform of the coursebook) makes me nervous.*" (M=1,98) item has the lowest average score. According to these findings, students had negative attitudes regarding speaking tests related to the reasons of the assessor/instructor. Paker and Höl (2012) also discussed that speaking examinations are the most difficult and stressful component of the language exams for students for a variety of reasons, one of which being the assessors' attitudes throughout the testing.

Although item 4 has the highest average score, item 2 "*I feel stressed over not making proper sentences and explaining my ideas in speaking exams*" (M=3,80), and item 8 "*Midterm exams make me nervous.*" (M=3,79) have quite close averages. These show that students are anxious during the speaking assessment due to linguistic problems and they also feel very anxious about mid-term exams. However; doing the tasks in the component is the least worrying thing for them.

When the mean values of the items in the enjoyment section were studied, it was determined that while an item from the speaking assessment section has the highest overall average, an item from the exams section has the lowest overall average. It was discovered

that "I24- I enjoy the speaking exam if the instructor's attitude is positive." (M=4,39) has the highest average score while "I26. I enjoy mid-term exams." (M=2,39) item has the lowest average score. As a result, the findings demonstrate that students have a positive attitude toward speaking tests when the assessor has a positive attitude as well. As previously said, assessors' attitude is one of the aspects that might influence students' attitudes during speaking tests (Paker and Höl, 2012), hence it is expected that students experience more joy as a result of the assessor's pleasant attitude. Furthermore, students believe the midterm exam to be the least enjoyable, which is consistent with the data from the anxiety section, which show that midterm exams are one of the most anxiety-inducing instruments.

Which of the assessment and evaluation instruments causes the highest and the lowest anxiety among A1-level students? The purpose of this question is to find which assessment and evaluation instruments create the most and least anxiety among A1-level students. Table 10 presents the means of the participants' total ratings from four assessment and evaluation instruments in the anxiety section of the questionnaire.

Table 10

Descriptive Statistics of Each Instrument in the Anxiety Section of the Questionnaire

Anxiety Items	Descriptive Statistics				
	N	Minimum	Maximum	Mean	Std. Deviation
Speaking Assess.	218	1,00	5,00	3,34	.75416
Exams	218	1,00	5,00	3,14	.98148
Writing Portfolio	218	1,00	5,00	2,86	.76617
Online Components	218	1,00	5,00	2,53	.87342

Table 10 shows that the mean of the students who supported the anxiety statements in the questionnaire indicative of speaking assessment is the highest, at 3,34, followed by the mean of anxiety items related to exams (M=3,14). This result is consistent with data obtained by Paker and Höl (2012). They also found that students had higher anxiety during the speaking test and that the speaking test was regarded as the most challenging test among the students when compared to the testing of other language skills in their study,

which aimed to explore the attitudes and perceptions of students and instructors towards the speaking test at a School of Foreign Languages. The exams in the current study have the second-highest mean in the anxiety section, and Trifoni and Shahini (2011) also found that a significant number of students were impacted, at least to some degree, by test anxiety.

Participants' responses to the anxiety items in the questionnaire about online component tasks revealed that it has the lowest mean ($M=2,53$), followed by the mean of anxiety items related to writing portfolio ($M=2,86$). Although the subjects' responses to these statements suggest that students could be anxious during the speaking assessment, they do not have strong negative emotions and attitudes about completing the tasks in the online component. When all of the different instruments' means are compared, the results show that exams ($M=3,14$) and speaking assessments ($M=3,34$) cause the most anxiety by leading to negative feelings with remarkable similar averages, whereas the online component and writing portfolio cause the least anxiety also with close averages $M=2,53$ and $M=2,86$, implying that students do not have highly negative attitudes toward them as they do not feel anxious or nervous about the components of those instruments or the requirements that those instruments entail.

Despite the fact that testing is a vehicle for determining scores, assessment and evaluation instruments in this study contain both score-oriented testing instruments that focus on the test result and process-oriented instruments that assist the learner in the learning process (Altay, 2007). As a result, it's comprehensible that speaking tests, quizzes, and midterm examinations with summative assessments generate greater anxiety, whereas the writing portfolio and online component with formative assessments cause less anxiety and are viewed as more likable by the participants. Another possible explanation is that students value teacher feedback and the continuous/gradual development of their writing skills (Deluca et al., 2017), and they regard the tasks in the online component as extra practice.

Which of the assessment and evaluation instruments causes the highest and the lowest enjoyment among A1-level students? The purpose of this question is to find which assessment and evaluation instruments create the most and least enjoyment among A1-level students. Table 11 presents the means of the participants' total ratings from four assessment and evaluation instruments in the enjoyment section of the questionnaire.

Table 11

Descriptive Statistics of Each Instrument in the Enjoyment Section of the Questionnaire

Enjoyment Items	Descriptive Statistics				
	N	Minimum	Maximum	Mean	Std. Deviation
Speaking Assess.	218	2,00	4,83	3,671	.56817
Exams	218	1,00	5,00	2,972	.76016
Writing Portfolio	218	1,00	5,00	3,676	.78860
Online Components	218	1,00	5,00	3,593	.80812

Table 11 reveals that the students who supported the enjoyment statements in the questionnaire indicative of writing portfolio and speaking assessment have nearly identical averages, at 3,676 and 3,671, respectively. The participants' responses to the enjoyment items in the questionnaire about exams revealed that they had the lowest mean ($M=2,97$). The subjects' responses to these statements suggest that students are relaxed and enjoy completing the tasks in the writing portfolio. Furthermore, the writing portfolio and the speaking assessment have strikingly similar means, indicating that students also had positive attitudes regarding the speaking assessment.

One possible explanation is that students appear to appreciate the challenge and are motivated to overcome it (Mierzwa, 2018) because speaking assessment is also an instrument that causes anxiety for them. As a consequence, speaking assessment has been able to strengthen enjoyment. Exams, on the other hand, appear to be the instrument that participants find the least enjoyable and have the least positive attitudes. As for writing portfolios, it's reasonable to assume that students see portfolios as a good learning and

evaluation tool because they allow teachers and students to cooperate on teaching, learning, and assessment (Azzarfam et al., 2016). Feedback from instructors, in particular, plays an important role in facilitating learners' performance on writing tasks and, as a result, enhancing their writing skills. As a result, it is viewed as an instrument that enhances enjoyment. Furthermore, the mean of students' responses to the items related to the online component in the enjoyment section ($M=3,59$) is fairly similar to the ones in the writing portfolio and the speaking assessment, implying that students have moderately favorable sentiments toward it.

Are there any significant differences in the attitudes of A1-level students regarding the extent of anxiety and enjoyment in terms of gender? This question attempts to determine whether there were any significant differences in the attitudes of A1-level students regarding the extent of anxiety and enjoyment in terms of gender within the scope of the research. The difference in mean scores obtained from the anxiety and enjoyment parts of the questionnaire according to gender was analyzed using an independent sample t-test. The results were reported based on the fact that the difference was statistically significant for variables with a p-value less than 0.05, but not for variables with a p-value more than 0.05. **Table 12** shows the descriptive statistics (mean scores and standard deviations) computed on the students' responses to the questionnaire's anxiety and enjoyment components.

Table 12

Independent T-test Results for Gender

Independent t-test results							
	Gender	N	Mean	SD	t	df	Sig.
Anxiety	female	104	3,20	.57244	3,057	216	0,003
	male	114	2,96	.60363			
Enjoyment	female	104	3,49	.43274	-0,321	216	0,749
	male	114	3,51	.47736			

In the anxiety items of the questionnaire, the statistical results showed that female participants' mean score ($M=3,20$) was higher than male participants' score ($M=2,96$). Table 12 shows that female students are more anxious than male students, implying that they have more unfavorable views about assessment and evaluation instruments than male students. When the p-value of 0,003 is compared to the level of significance, which is 0,05, it is seen that our value is less than the defined significance level of 5%, indicating that the mean score between male and female students is significantly different.

Dalkılıç (2001) investigated freshmen ELT students using quantitative and qualitative methods and discovered that female ELT students are significantly more anxious than male ELT students. Pappamihel (2002) found that female students are more anxious than male students in a study of middle-school Mexican immigrant students in the United States. Trifoni and Shahini (2011) also found that test anxiety is associated with the gender of the subjects and female students feel more worried than males. Lastly, Demirdağ (2012) discovered that female students' mean anxiety score was higher than male students, and the difference was highly significant, implying that female students are significantly more anxious than male students, in his research on the relationship between foreign language anxiety and language performance. Aida (1994), Chang (1996), Onwuegbuzie et al. (1999), and Özer and Altay (2021), on the other hand, found no significant relationship between gender and foreign language anxiety.

On the other hand, the statistical results for the questionnaire's enjoyment items revealed that female participants scored 3,49 while male participants scored 3,51, indicating that male participants scored slightly higher than female participants. When the p-value of 0,749 is compared to the threshold of significance of 0,05, it can be observed that our value is more than the set significance level of 5%, suggesting that the mean score between the groups is not significantly different. This finding is consistent with Özer and Altay's research (2021), which indicated that gender differences had no effect on participants' enjoyment levels, and Mierzwa's research (2018), which aims to investigate the influence of foreign

language enjoyment on learning English as a foreign language and the relationship from the perspective of gender. In other words, there was no significant difference in the enjoyment of language learning between males and females in their studies.

Are there any significant differences in the attitudes of A1-level students regarding the extent of anxiety and enjoyment in terms of academic major? Within the framework of the research, this question attempts to determine whether there were any significant differences in the attitudes of A1-level students regarding the extent of anxiety and enjoyment in terms of academic major. Academic majors are grouped into two categories based on the faculties where the participants will study: Social Sciences and Natural and Applied Sciences. The students' majors that take place in the scope of Social Sciences included departments at the Faculty of Education (31,8%), Faculty of Economic and Administrative Sciences (13,8 %), Faculty of Arts and Sciences (3,8 %) while the majors that take place in the scope of Natural and Applied Sciences included the departments at the Faculty of Architecture (1,8 %) and Faculty of Engineering (42,7 %). An independent sample t-test was used to examine the difference in mean scores obtained from the anxiety and enjoyment sections of the questionnaire according to academic major. The difference was statistically significant for variables with a p-value less than 0.05, but not for those with a p-value more than 0.05, hence the findings were reported.

Table 13

Independent T-test Results for Academic Major

Independent t-test results							
	Academic Major	N	Mean	SD	t	df	Sig.
Anxiety	Social Sciences	121	3,18	.58783	2,995	216	0,003
	Natural and Applied Sciences	97	2,94	.59149			
Enjoyment	Social Sciences	121	3,39	.42557	-4,032	216	0,001
	Natural and Applied Sciences	97	3,63	.45840			

The statistical results showed that the mean score of participants who will major in Social Sciences ($M=3,18$) was higher than the mean score of participants who will major in Natural and Applied Sciences ($M=2,94$) on the anxiety items of the questionnaire. **Table 13** reveals that students choosing to major in Social Sciences are more anxious than those going to major in Natural and Applied Sciences, meaning that those choosing to major in Social Sciences have more negative attitudes regarding assessment and evaluation instruments than the remaining portion. When the p-value of 0,003 is compared to the level of significance, which is 0,05, it can be observed that our value is less than the predetermined significance level of 5%, showing that the mean score between Social Sciences students and Natural Sciences students differs significantly.

The statistical results in Table 13 also showed that the mean score of participants who will major in Natural and Applied Sciences ($M=3,63$) was higher than the mean score of participants who will major in Social Sciences ($M=3,39$) on the enjoyment items of the questionnaire. As shown in Table 13, students who are going to major in Natural and Applied Sciences enjoy various components and requirements of the testing instruments and have more positive attitudes about assessment and evaluation instruments than students who are going to major in Social Sciences. When the p-value of 0,001 is compared to the threshold of significance, which is 0,05, it is clear that our value is significantly lower than the established significance level of 5%, showing that the mean score between students in Social Sciences and students in Natural and Applied Sciences is significantly different.

Taken together, when compared to students to be major in Natural Sciences, it appears that students to be major in Social Sciences are more anxious and enjoy less when taking tests (i.e. quizzes, midterm exams, and speaking exams) and completing activities in the writing portfolio and online component, which indicates that while students to be major in Social Sciences have negative attitudes toward assessment and evaluation instruments, students to be major in Natural Sciences have more positive attitudes.

Higher education's overarching purpose is to assist students in discovering and developing their full potential by challenging and encouraging them to integrate it comprehensively (Sari & Arsyad, 2021). Even though teaching is the primary goal of education, testing is an essential component, as teaching cannot be considered without assessing the effectiveness of the teaching-learning environment (Altay, 2007). As a result, a variety of assessment and evaluation instruments are employed to measure students' various language skills. However, considering it is their first year at university, this may have been a challenge for students, particularly in the prep class. Sari and Arsydad (2021) discovered that there is a significant difference in the ability to adjust between students in science and social studies study programs. They revealed that science study students find it easier to adjust to new or hard situations since they have a more critical, objective, and open attitude. Given that testing is typically an unpleasant and stressful experience for students, that mindset might have been useful in reducing anxiety among Natural Sciences students.

Qualitative Data Analysis

Findings From the Student Interviews

This section presents the results from the interviews. The researcher conducted semi-structured interviews with volunteer A1-level students after preparing questions about what students are anxious about and what factors they enjoy regarding four distinct assessment and evaluation instruments used in the department. The semi-structured interviews were conducted in Turkish so that the participants could express themselves more freely, taking into account the fact that the students are A1-level.

The descriptive analysis method and thematic analysis were used to analyze the data collected from the students. The following steps are followed:

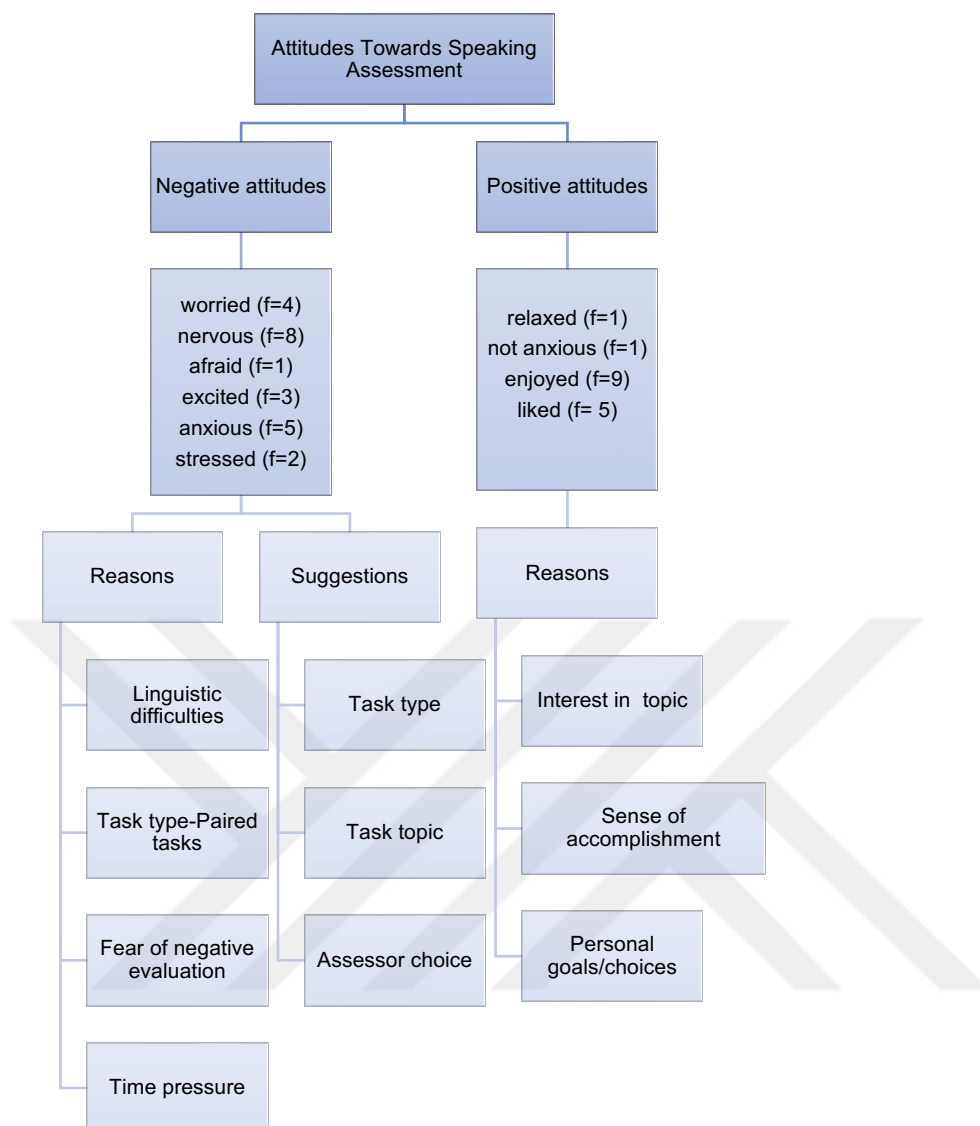
- The themes of positive and negative attitudes recurred throughout the dataset, so students' attitudes were categorized as positive and negative attitudes based on the codes that emerged from the interviews.
- These recurrent themes are also divided into three sub-themes named "reasons for negative attitudes", "suggestions that can ease negative attitudes" and "the reasons for positive attitudes".
- The data under each category is read and gone through line by line to code as much as possible.
- Words and phrases that are recurring and important in representing the themes are coded, grouped, and labeled. These labels are usually words and short phrases.
- To illustrate how they relate to one another, a figure was created.

What factors emerge from speaking exams that affect participants' attitudes in the context of anxiety and enjoyment? This question was designed to gather information on A1-level students' attitudes toward speaking assessments and to highlight the factors that can influence these attitudes in terms of anxiety and enjoyment.

The analysis of the qualitative data showed that half of the interviewees felt that speaking exams could be a source of stress and a situation that could create anxiety for them. When they were asked whether there are some sides they enjoyed, they responded positively by explaining their reasons except for one interviewee. Figure 5 below depicts the model for both positive and negative attitudes as well as possible solutions for negative attitudes.

Figure 5

Attitudes Towards Speaking Assessment



The qualitative data analysis revealed that half of the participants agreed that speaking assessments are a source of anxiety for students for a variety of reasons. Moreover, even if the other half stated that they felt less nervous or relaxed during the speaking exams, they also mentioned some factors that could cause anxiety for them. When discussing their feelings about the speaking assessment, interviewees chose to use the words shown in the figure under the negative attitudes section. For example, some interviewees commented:

“I feel worried about the speaking exam” (P1)

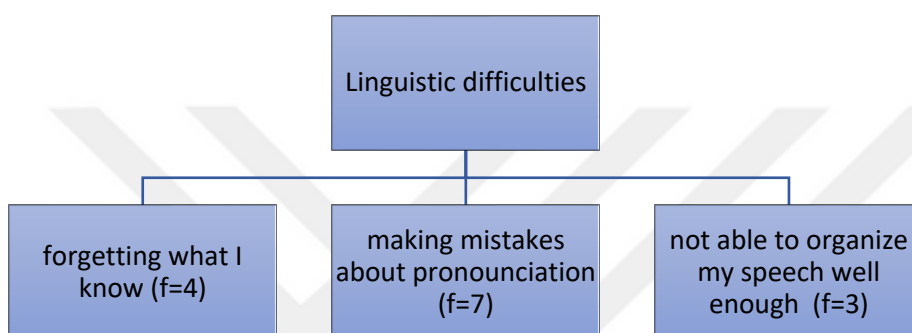
“I feel anxious because I’m having serious difficulties in the speaking exam” (P2)

“At first, I feel nervous about performing poorly, but then I feel all right” (P5)

The data also indicated that all participants had some reasons for having negative attitudes and feeling anxious or nervous. Figure 6 demonstrates the codes and frequencies regarding linguistic difficulties:

Figure 6

Linguistic Difficulties



As is seen in Figure 6, some of the reasons are categorized under the label of linguistic difficulties as the participants stated that they sometimes forgot some structures and vocabulary items and had difficulties in organizing what they were going to say. Moreover, some of them expressed that they were anxious about pronouncing the words wrong. Some of the comments regarding linguistic difficulties were presented below:

“I’m nervous about not organizing my speech well and making pronunciation mistakes”(P8)

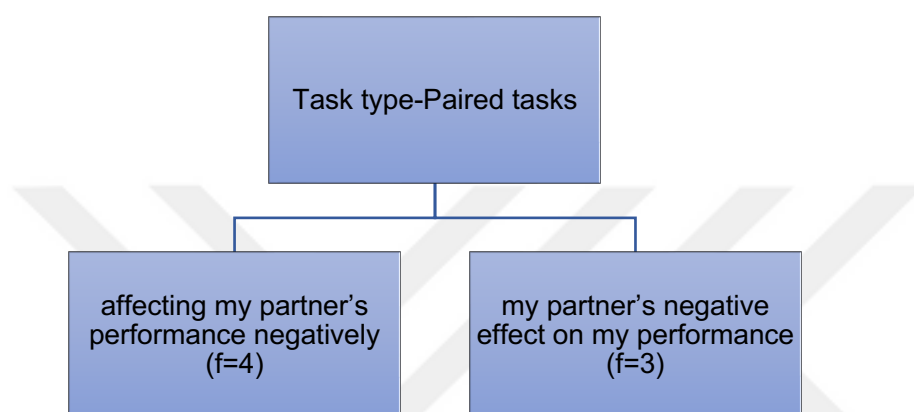
“The thing that makes me the most anxious is that making pronunciation mistakes (P2)

“Making grammatical mistakes and pronouncing words wrong make me anxious” (P9)

Another factor that emerged from the interviews and caused negative attitudes is how the task is performed. In other words, task type- pair work, group work, or individual work- is one of the elements that can affect how students think and feel about the speaking exam. Figure 7 shows the details of the “Task type-Paired tasks” category:

Figure 7

Task Type- Paired Tasks



The participants pointed out that paired tasks in the speaking exams caused anxiety for them as they were worried about affecting their partners' performance negatively or being affected negatively by their partner's possible poor performance. The following comments were made:

“We perform the tasks with a lot of discomfort in the speaking exam and our poor performance can cause problems and affect our partner's performance negatively and this situation makes me very nervous.” (P3)

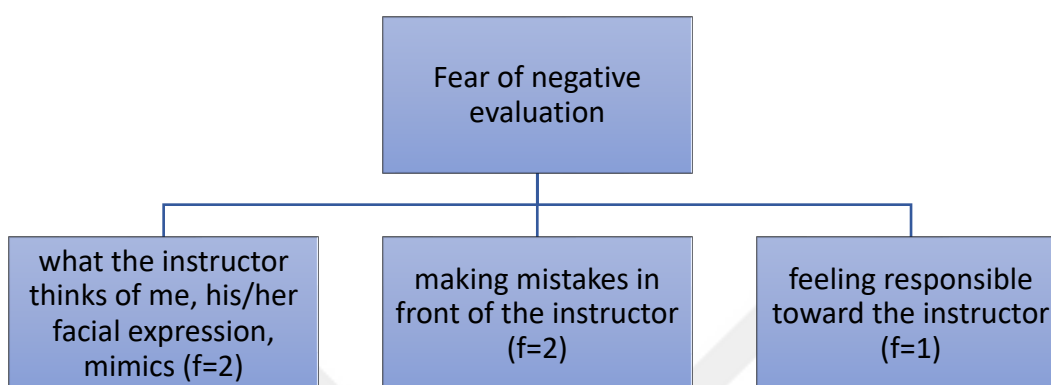
“The previous speaking exam was performed with a pair and I think if my partner is a worried person or cannot talk due to being anxious, it can affect my performance in a negative way and this is something that makes me nervous.” (P4)

Some of the participant students also indicated that the fear of negative evaluation is another reason that makes them anxious or nervous in the speaking exams. Since the instructor is the one who teaches them, they usually mentioned that they were anxious or

worried about what the instructor would think about their performance. Figure 8 shows the details of this category.

Figure 8

Fear of Negative Evaluation



Some of the participants highlighted the fear of negative evaluation by the instructor as a factor making them feel uneasy or nervous during speaking exams. The comments below were made:

"One of the main sources of anxiety for me during speaking tests is what the instructor thinks of me because I believe I am not very good at English. The instructor is the person who teaches us. I'd be disappointed if I taught someone and they didn't perform well. That is why the instructor is working so hard to teach us, and it makes me nervous about my speaking exam performance." (P1)

"... sometimes I worry about making mistakes in front of the instructor." (P5)

"When I speak English in the exam, I feel anxious about making a mistake in front of the instructor." (P8)

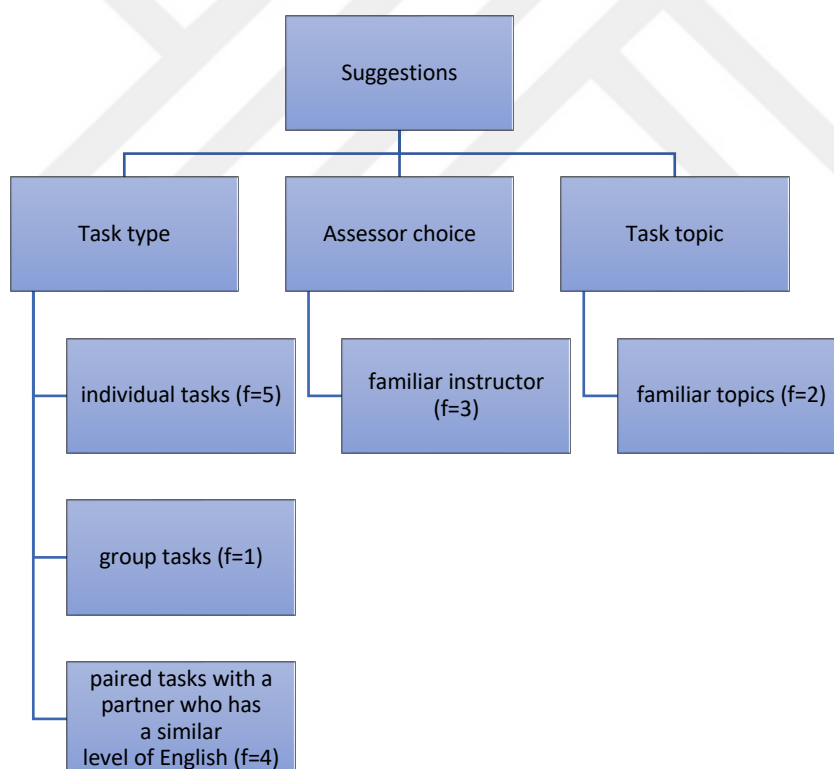
For some participant students, time pressure is another reason for feeling anxious during speaking exams. For this category, one participant mentioned that she felt nervous in the speaking exams since they were expected to speak for a specific time. The comment below was made:

“We're also required to speak for a particular amount of time, perhaps 2-3 minutes. Despite my preparation, I occasionally worry about my ability to speak effectively for the required period of time..” (P9)

The data analysis also reveals that the participants have some suggestions for overcoming or reducing the effects of these negative reactions. Participants suggested using a different task type, taking the speaking test with a familiar instructor (this refers to the teacher in that classroom), and having the exam include themes that have already been covered in-class activities can help them to feel less anxious. This category's details are shown in Figure 9.

Figure 9

Suggestions for Reducing the Effects of Negative Attitudes Toward Speaking Assessment



Paired tasks were noted as a source of anxiety for students in the previous section. As a result, participant students stated that taking the exam alone- without a partner or with a partner with a similar English level to theirs- made them feel less anxious and nervous. They also stated that taking the exam alongside their classroom teacher and encountering

common speaking subjects can assist them in developing a more positive outlook. Some comments of them are:

“Having an instructor I recognize in the exam helps to reduce my anxiousness. I’ll be nervous if I take the exam with a teacher I don’t know.” (P1)

“I suppose I’d feel more at ease if I took the exam alone because I’d feel more at ease with the teacher one-on-one.” (P2)

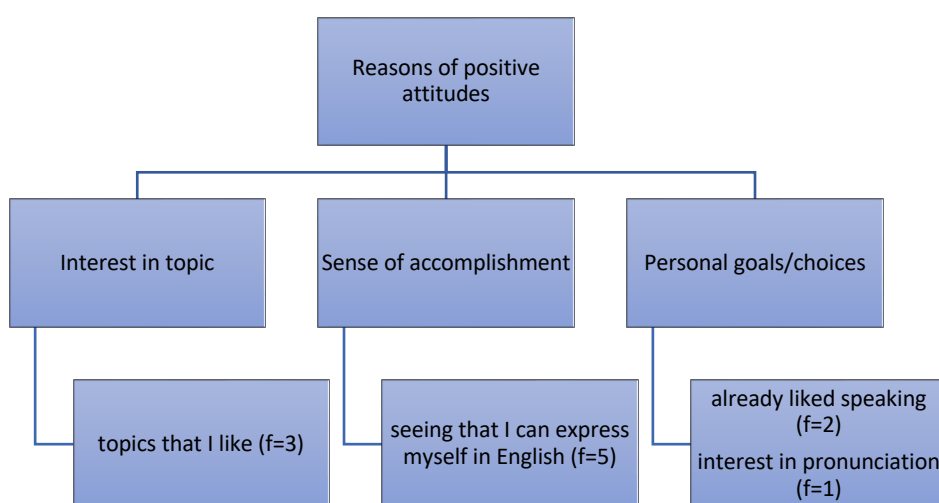
“Also, I believe that partnering with someone whose English level in the exam was similar to mine would reduce my stress level.” (P5)

“Being paired with a partner of a similar level on the exam makes me feel good since realizing that he or she is just like me reduces my anxiety. I suppose I’d be worried if my speaking partner was particularly good or particularly poor.” (P8)

The qualitative data analysis revealed that some of the participants have positive attitudes concerning speaking assessment and particularly enjoy some aspects of it because of some reasons. Figure 10 indicates the details of this sub-theme.

Figure 10

Reasons for Positive Attitudes Toward Speaking Assessment



According to the data, some of the participants enjoy speaking exams when they come across issues that interest them. Some of them stated that they enjoyed the speaking exam since it allowed them to put their speaking skills to the test and see how well they could express themselves in English. In other words, they like the feeling of accomplishment they get from speaking. Furthermore, some of them indicated that the speaking evaluation was enjoyable because they already enjoyed speaking or had a particular interest in pronunciation. The following comments were made:

“When the subject of the speaking test is something I like or is relevant to my field of interest, it brings me joy.” (P2)

“The speaking test reflects our actual performance. To put it another way, we see what we can and can't do, whether we're making progress or not, which is why I enjoy it.” (P5)

“Speaking exams appeal to me since I enjoy speaking and am also interested in pronunciation.” (P10)

To conclude, half of the participants had unfavorable views concerning speaking assessment, whereas the other half have positive sentiments. In the scope of the speaking assessment, it can be suggested that some of the participants experience communication anxiety. Communication apprehension, or the fear of communicating with others, can manifest itself in situations when learners lack communication skills or lack the requisite linguistic expertise to express their ideas and views (Gardner & McIntyre, 1993).

To begin with, the most common concern is pronunciation; the majority of participants are concerned about pronouncing the words incorrectly and they also worry about not performing well due to linguistic difficulties, which produces unfavorable attitudes toward the speaking exam. Because of their poor skill level in the target language, some of the participants in Bata and Castro's study (2021) who showed signs of anxiety, nervousness, and worry made more speech mistakes in speaking assessments. According

to the findings of the current study, a lack of linguistic knowledge or a low-performance level can cause anxiety and negative attitudes about the speaking assessment.

In Wörde's (2003) study which aims to reveal students' perspectives on foreign language anxiety, the participants also cited numerous and various sources for their anxiety, speaking activities, fear of negative evaluation, and the instructors themselves were sources of anxiety reported by the participants. Participants in the current study were also concerned about what the teacher thought of them, which is an indication of the negative evaluation fear that occurs when foreign language learners believe they are unable to make the best impression possible about their performance (Horwitz et al. 1986).

Students also report feeling overwhelmed and anxious while speaking, possibly as a result of a restricted vocabulary or grammatical knowledge. Paker and Höl (2012) conducted a study at a School of Foreign Languages to investigate the attitudes and perceptions of students and instructors toward the speaking test. They discovered that students had higher anxiety during the speaking test and that they were unable to express themselves adequately. Thus, the results of the current study are consistent with the abovementioned studies.

Another important finding from the interviews is that paired assignments induce anxiety in students because they are afraid of negatively affecting their partner's performance or vice versa. Factors such as teamwork and time constraints in the speaking exam were found to generate nervousness and anxiousness in Bata and Castro's study (2021), as students did a role play in which they had to ask and answer questions from their partners within a time restriction. They also discovered that taking the speaking exam in pairs had a negative impact on the partners' performance, as some participants stated that they didn't feel at ease with their partners who appeared anxious during the exam, which corresponds to the findings of this study. Finally, students dislike speaking exams because they are afraid of being negatively evaluated by the instructor.

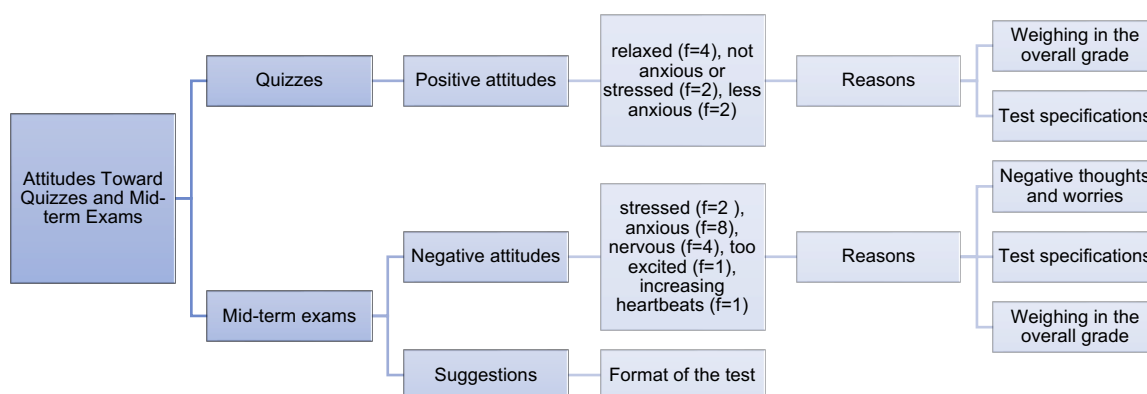
In terms of recommendations for feeling more confident in speaking exams, participants said they prefer to take the exam alone or with a partner who has a comparable English level to them. Furthermore, the fact that the class instructor is the examiner for the speaking exam is a soothing aspect for them. Half of the participants, on the other hand, believe that speaking exams are favorable, and the most prominent factor that emerged from the interviews regarding enjoyment of speaking assessments is that students can feel accomplished, and they regard their performance in speaking exams as concrete proof of their development.

What factors emerge from quizzes and mid-term exams that affect participants' attitudes in the context of anxiety and enjoyment? This question was intended to obtain information on the attitudes of A1-level students regarding quizzes and mid-term exams, as well as the elements that can affect these attitudes in terms of anxiety and enjoyment.

The qualitative data revealed that, with the exception of two students, practically everyone has a favorable attitude regarding quizzes, while they have a negative attitude concerning mid-term exams. They also mentioned the factors that influence their attitudes. The hierarchical structure shown in Figure 11 highlights the reasons for positive attitudes regarding quizzes, negative attitudes toward mid-term examinations, and potential solutions for negative attitudes.

Figure 11

Attitudes Towards Quizzes and Mid-term Exams



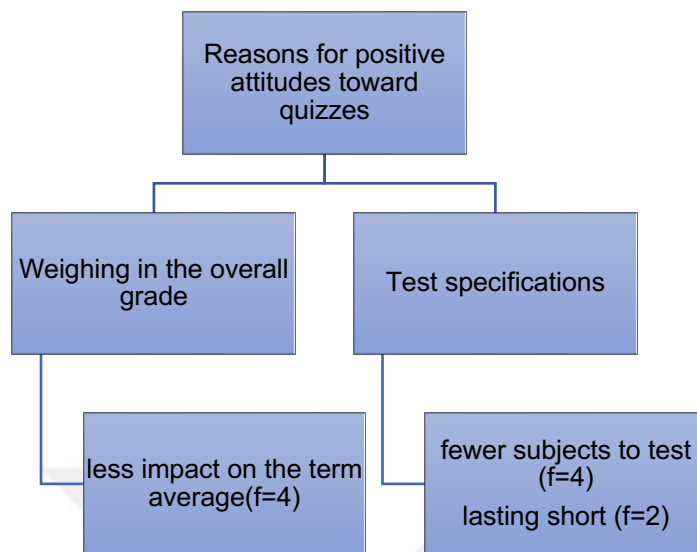
Despite having different attitudes toward quizzes and mid-term exams, all of the participants (f=9) agreed that quizzes and mid-term examinations were valuable because they allowed them to test themselves and see if they could do well or not. Furthermore, they like receiving feedback after these exams because they could identify the parts they struggled with. They value the ability to recognize their weaknesses through quizzes and mid-term exams. As for the content of it, they made the following comments:

"I enjoy tests because they allow me to see my flaws in post-exam feedback and learn from them...I believe that when we make mistakes and receive feedback, we learn the correct version of our faults and it sticks with us longer." (P4)

"Exams help me realize my flaws and drive me to study, which I enjoy." (P7)

"Exams are something I enjoy. I believe that quizzes and midterm exams assist us in developing a regular study habit...In the post-exam feedback phase, I can notice my weaknesses, so I take notes and use them to improve.." (P10)

Despite the fact that quizzes and mid-term examinations are both types of formal assessment, the data analysis reveals that attitudes toward them differ. According to the participants, nearly all of them share a positive attitude that includes feeling relaxed and less nervous, based on their word choices to convey their ideas. Figure 12 shows the details of this category.

Figure 12*Reasons for Positive Attitudes Toward Quizzes*

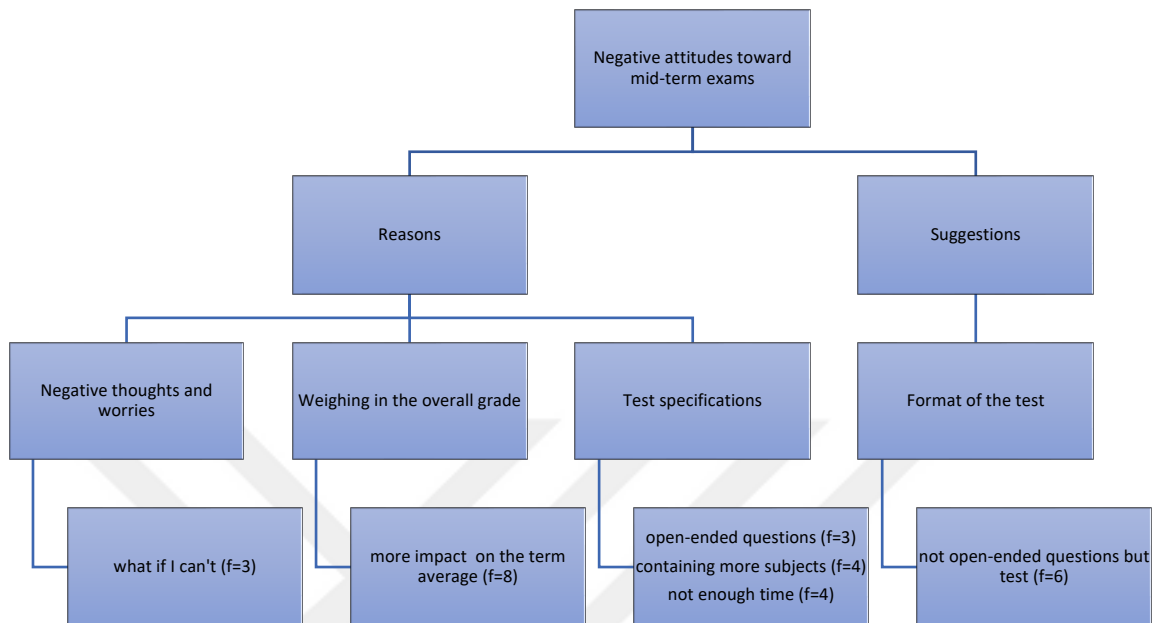
Participants said they felt more relaxed about quizzes because the proportion of quizzes makes up a lesser percentage of the overall term average, quizzes are shorter, and they have fewer subjects to study for the quizzes. Some comments are:

"I don't get anxious and nervous over quizzes because they don't affect our term average much." (P7)

"In fact, I don't have this sense (of being stressed) during quizzes because we can be better prepared because quizzes are narrower in scope." (P8)

"Because quizzes have a more limited scope, I feel more at ease and worry less. To put it another way, we learn the subjects and take the exams from the parts that we study." (P9)

On the other hand, almost all participants stated that they felt nervous, anxious, or stressed during midterm examinations, therefore they mostly have a negative attitude about them for a variety of reasons. Because the participants have some thoughts about why they have this attitude, they respectfully provide a proposal that they hope will make them feel better. The specifics of this category are shown in Figure 13.

Figure 13*The Reasons for Negative Attitudes Toward Mid-term Exams and Suggestions*

The weighting of midterm tests in the total grade was one of the most common notions that caused negative attitudes among participants. Participants are more anxious during mid-term exams because they make up a larger proportion of the term average. They also listed certain elements that can negatively affect individuals, such as the exam format, test scope, and time allotment, as well as some negative ideas that may arise from low confidence. The comments made by the participants were presented below:

“Midterms make me nervous since they are more significant to me. Worries about not being able to accomplish it or doing it badly cross my mind. Midterms give me anxiety because they are detailed and have a large impact on the average.” (P2)

“Midterm examinations have a greater impact on our average and taking the exam knowing this fact adds to the stress. I have some thoughts like “I can’t do it” during midterm exams.” (P8)

“Midterm examinations give me a lot of anxiety because the topics are becoming more comprehensive, and I’m afraid of messing up. It also has a greater impact on our semester average than the quiz. There are also open-ended questions, which are more concerning to me.” (P9)

According to the points above, midterm exams involve open-ended questions, which causes students anxiety. As a result, the majority of participants felt that a multiple-choice exam would reduce their anxiety. The comments below were made:

“I think midterm examinations with multiple choice questions rather than open-ended questions lessen my anxiety.” (P1)

“Perhaps I will be less stressed if the midterm exam is a test. When the exam is open-ended, it takes longer because I have to consider whether or not what I wrote is correct, which wastes time. When the exam is a test, though, I feel more at ease because I have a choice in front of me.” (P2)

“If the midterm exam were a test, I believe I would be less concerned because choosing from the options when taking a test makes my job easier and reduces the chances of making mistakes.” (P9)

To summarize, the findings of the qualitative study revealed that participants have a positive attitude toward quizzes but a negative attitude toward mid-term exams. Thus, it's possible that students experience test anxiety in midterm exams, which influences their attitudes regarding this instrument. This result is consistent with Aida's (1994) findings, which revealed that test anxiety and fear of failure were two factors that influenced the participants' anxiety-related ideas in her study on the relationship between language anxiety and Japanese language learning.

Despite their differing perspectives on quizzes and mid-term exams, all of the participants thought that they were impactful since they allowed them to put themselves to the test and see if they could do well enough. One of the most prominent factors that

emerged from the data is the weighting of quizzes and mid-term exams in the overall grade of the students. Participants perceive this factor as positive for quizzes because of the lower impact, but as negative for mid-term exams, because it affects the term average more. Allen (2005) stated that students' exam grades are one of the most reliable indicators of their academic progress since they reflect if they are mastering the academic knowledge they have learned. Because grades form part of a student's permanent record, they are likely to have a significant impact on how they feel. It is anticipatable that the importance of these grades is regarded by students and has an impact on their attitudes.

Furthermore, in terms of quizzes and mid-term examinations, factors such as test scope and time allocation that appear in the test specifications label are evaluated differently by participants. In other words, although a narrower test scope and shorter exam time in quizzes are factors that lead to favorable attitudes, broader test scope and relatively insufficient exam length in mid-term exams lead to negative attitudes. As a result, the majority of participants stated that if the exam style is multiple-choice, they will have a more positive attitude regarding exams since they believe it will lessen their negative thoughts and concerns. This is in line with Struyven's (2005) findings, which found that the multiple-choice format is preferable to alternative formats due to participants' expectations of less anxiety and complexity, a lower level of difficulty, and a better rate of success.

What factors emerge from the writing portfolio that affects participants' attitudes in the context of anxiety and enjoyment? This question was designed to learn about A1-level students' attitudes toward the writing portfolio, as well as the factors that can influence these attitudes in terms of anxiety and enjoyment.

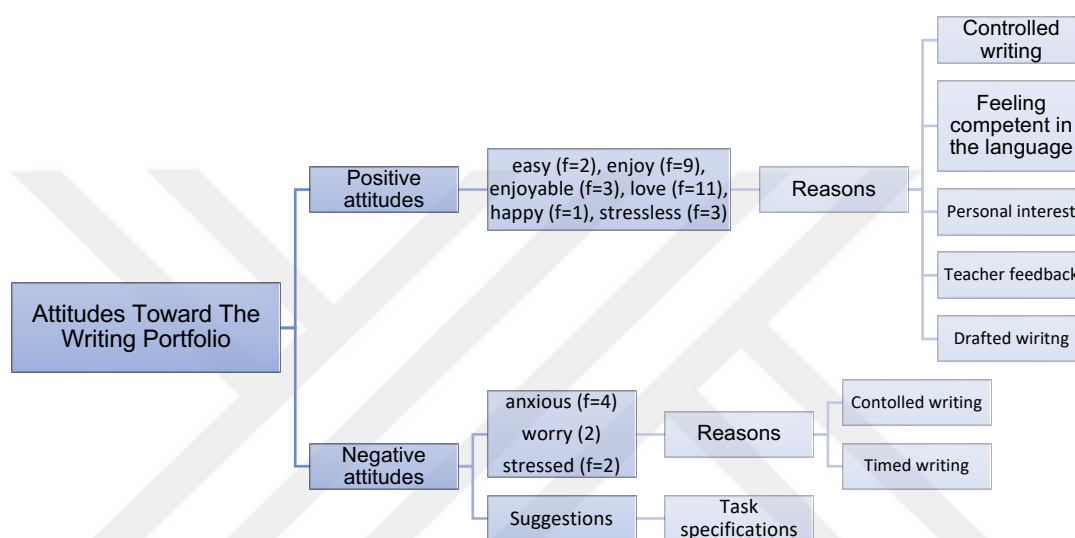
Excluding one student, the qualitative data suggested that almost everyone had a positive opinion about the writing portfolio. They also discussed the influences on their attitudes. In the context of this, some comments can be found below, and figure 14 depicts the reasons for favorable and negative attitudes toward the writing portfolio, as well as suggested solutions for negative perspectives.

"I enjoy completing the assignments in the writing portfolio...I have no worries about this." (P2)

"In the writing portfolio, I don't sense anything that makes me nervous. I'm not bothered by anything." (P8)

Figure 14

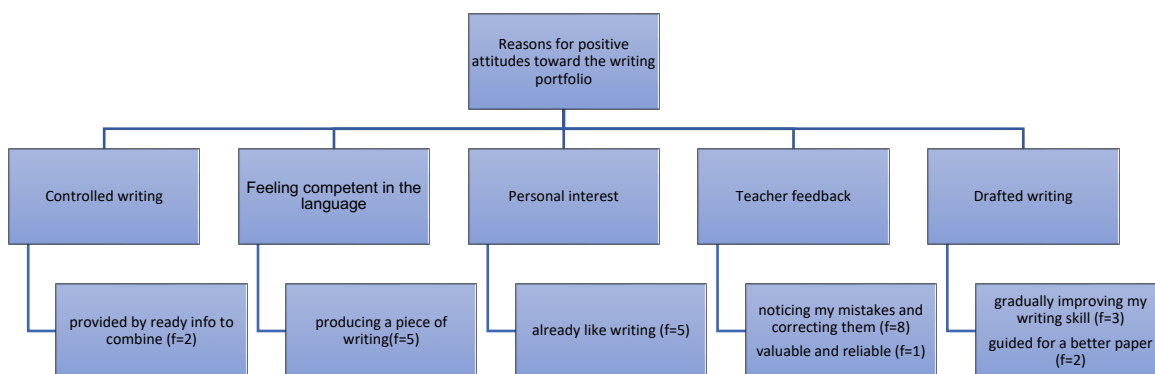
Attitudes Toward the Writing Portfolio



Based on what the participants reported, the data analysis demonstrates that there are a range of reasons for having positive opinions toward the writing portfolio. Figure 15 shows the details of this part, along with the frequencies.

Figure 15

Reasons for Positive Attitudes Toward the Writing Portfolio



As can be seen in figure 15, the most frequently stated reason is that they love and value instructor feedback (f=8) on their product since it allows them to identify and correct mistakes. Participants also mentioned that writing a second draft allows them to gradually improve and guides them in producing a higher-quality paper based on the comments they receive. Some of the comments were presented below:

"I like getting feedback because it allows me to see my faults and correct them; even if I make a lot of mistakes, I don't mind the teacher's correction because learning is more important to me. Writing the second draft is also enjoyable since a more successful piece develops." (P5)

"I enjoy receiving feedback because it makes me happy rather than worried because it allows me to correct my mistakes." (P6)

"I also enjoy that our writing portfolio requires two drafts of our work so that we can identify our flaws and correct them...I also appreciate that our teacher provides feedback because this individual is a qualified expert, and I believe it is more valuable and reliable than feedback from any writing website." (P7)

Another aspect that emerged was that they enjoy producing written work (f=5), which is a measure of their English proficiency since being able to generate written work using vocabulary or grammatical structures that they have acquired so far is important to them. Furthermore, according to interview data, some participants have a positive attitude toward the writing portfolio because they already enjoy (f=5) writing in their mother tongue. Lastly, they also mentioned that they felt at ease since the tasks in the portfolio are controlled. The following comments were made:

"It's enjoyable for me since it allows me to produce something with the vocabulary and grammatical structures I've learned." (P5)

"I have positive emotions because I feel that I am producing something rather than being measured....Because I love writing, I also enjoy completing the tasks in the writing portfolio." (P7)

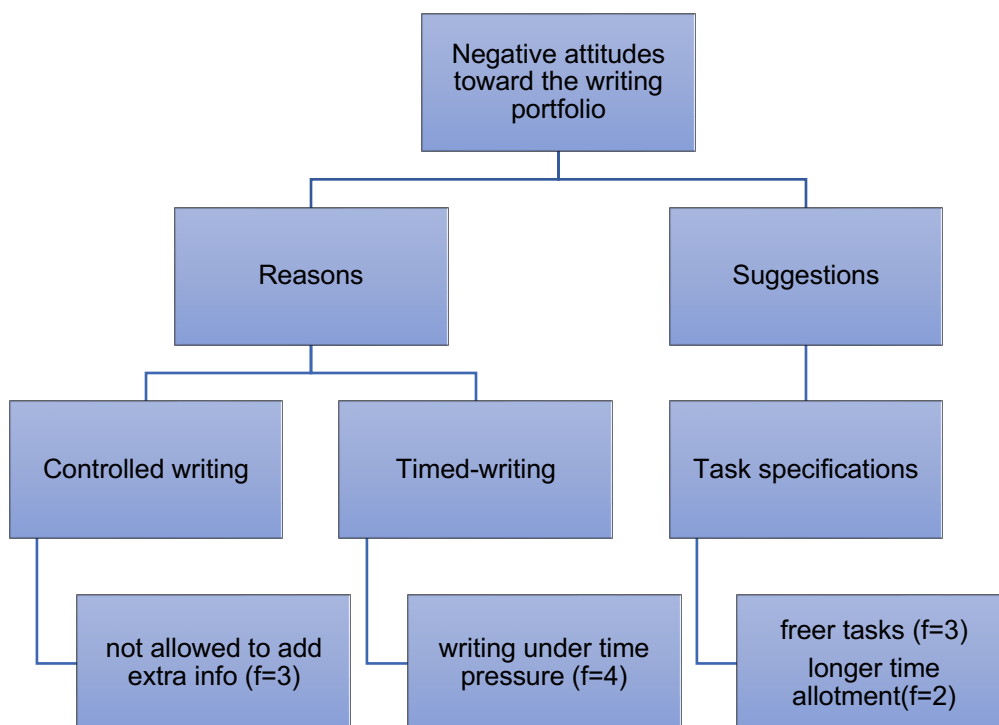
"I enjoy writing, therefore the writing portfolio appeals to me." (P4)

"That some information is provided us to use for the tasks the writing portfolio makes it simple for me to write and allows me to complete the task without difficulty; in fact, I enjoy it." (P1)

Except for one, none of the participants believe the writing portfolio is an anxiety-inducing assessment tool for them. However, two aspects of the writing portfolio were identified by the participants as causing anxiety or being stressful for them. Figure 16 shows the reasons for unfavorable attitudes about the writing portfolio, as well as ideas for strengthening positive attitudes.

Figure 16

The Reasons for Negative Attitudes Toward the Writing Portfolio and Suggestions



As shown in figure 16, writing under time limitations and controlled activities, in which the information is already supplied to the students and they are asked to synthesize or organize it, were recognized by the participants as two parts of the writing portfolio that cause anxiety or stress. As a result, they indicated that more free tasks and more time for the writing portfolio task can help them feel less anxious and more positive about the writing portfolio. The following comments were made:

"I believe I would appreciate it better if the writing portfolio tasks were less controlled and freer. In other words, we are given all of the information and then integrate it to produce sentences; this is not enjoyable...Writing in a short amount of time stresses me out." (P4)

"We are frequently provided controlled tasks in the writing portfolio, and adding extra content is not welcomed or encouraged. I'm a little worried about this. I think the tasks should be freer." (P1)

"Writing under time constraints might be difficult at times because we're desperately attempting to finish the paragraph. I believe I would be less anxious if the time allotted to do the assignment was extended." (P10)

To summarize, the qualitative analysis revealed that all of the participants, with the exception of one, have a distinctly positive attitude about the writing portfolio. Receiving feedback from the teacher is one of the most obvious explanations for this positive attitude that came from the data. According to the students' statements, they place a high value on teacher feedback since they perceive it as a vital source that can guide them and point out their mistakes.

Some students have a positive attitude toward the writing portfolio for personal reasons such as a special passion or enthusiasm for writing. Writing is important to the participants because they believe it is a sign of their language ability and progress. This result mirrors those of the previous studies that have examined the portfolio assessment.

Slater (1996), Segers and Dochy (2001), and Struyven (2005) found that students appreciate portfolio assessments because they believe they can internalize the material while working with it, which fosters deep-level and ongoing learning. To put it another way, the researchers asserted that students appreciated the time they spent building portfolios and that the writing portfolio, as a tool for progress and improvement, aided their learning.

On the other hand, despite the fact that virtually all of them hold overwhelmingly favorable views, they have discussed various concerns that may jeopardize their optimism. While some of the participants dislike writing relatively controlled tasks, others are stressed because the writing portfolio places them in a scenario where they must write under time constraints. As a result, they proposed that free tasks be included in the writing portfolio more frequently, and the time allotment be increased to decrease stress.

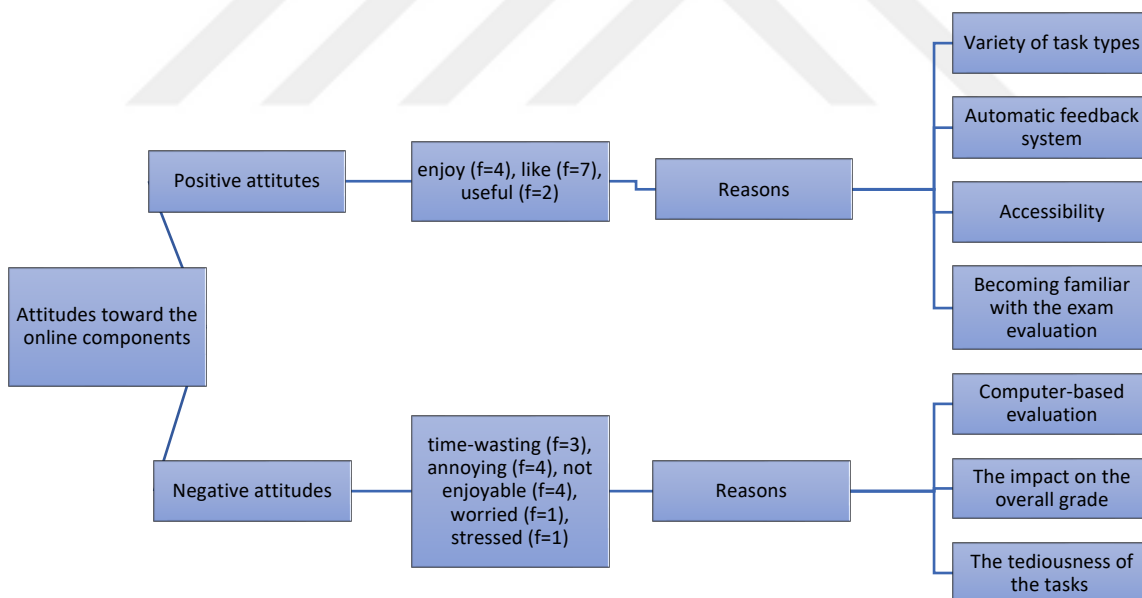
When writing is utilized as a vehicle for language development, McDonough and Fuentes (2015) noted that it is critical to give activities that encourage L2 writers to use their linguistic resources. Based on student interviews, it appears that the type of writing activity had a notable impact on students' attitudes, with participants expressing a preference for more free-form activities in which they could write their original thoughts using their linguistic knowledge. The writing portfolio tasks which are one of the foci of this study, on the other hand, provided a large set of prompts that students were expected to use, which they saw as restrictive and dissatisfying. McDonough and Fuentes (2015) argued that writing task prompts may be required for contexts where writing is used to assess students' language use, notwithstanding students' preferences. Even if some participants see controlled tasks as a source of negative feelings, it appears that they serve their purpose when the participants' linguistic knowledge and level (A1) are taken into account. Furthermore, participants in McDonough and Fuentes' study (2015) also expressed "timing concerns" while considering their options for collaborative and individual writing. Although the focus of this study is not on the task condition, it may be argued that time constraints are an important consideration for students in both studies.

What factors emerge from the online components that affect participants' attitudes in the context of anxiety and enjoyment? This question was formed to learn about A1-level students' attitudes toward the online components, as well as the aspects that can influence their anxiety and enjoyment levels.

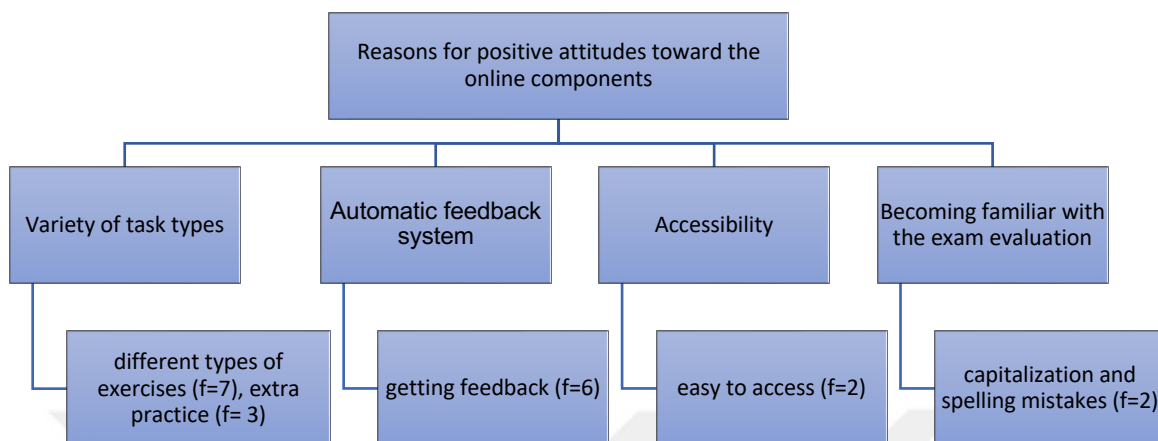
The qualitative data indicated that the majority of the participants had a lukewarm reception to the online components, with two statings, "There is nothing I like or favor about the online component." Even if the majority of the participants have favorable opinions on the online components and provide reasons for their perspectives, they all discussed certain aspects that they did not enjoy or appreciate. The factors for positive and negative attitudes regarding the online components are depicted in Figure 17.

Figure 17

Attitudes Toward the Online Components



The data analysis shows that there are a variety of reasons for having pleasant perceptions about the online component, based on what the participants said. This part's details, as well as the frequencies, are shown in Figure 18.

Figure 18*Reasons for Positive Attitudes Toward the Online Components*

The most often mentioned reason, as seen in Figure 18, is that they can complete a range of tasks, and the platform gives them more and varied possibilities for practice. The following comments were made:

"To begin with, I can state that I enjoy doing the activities in the online component because they are varied. There aren't the same types of questions every time. Listening, writing, filling in the blanks, and correcting mistakes are all alternatives."
(P2)

"I like the online exercises since they allow us to engage in a variety of activities. Some general tests and tests repeat a certain subject, which I like because it provides variety." (P1)

"I like it as it provides me with additional practice. I don't have a lot of resources, thus this platform is convenient for me." (P9)

In addition, the value of receiving feedback on the activities they had performed online was frequently mentioned by the participants. Before revealing the answer key, this automatic feedback method lets them recognize their mistakes and correct them. The

influence of an automatic feedback system on students appears to be positive, as they believed and indicated that they can see their mistakes and have several opportunities to correct them rather than only learning the correct answer. They made the following comments about the content:

“The nice thing is that if I make a mistake in the exercises, I can quickly correct it because I can receive immediate feedback.” (P1)

‘For each activity, I receive feedback. But I don't look at the correct answer straight away; instead, I attempt to find the correct answer first, then look at the correct version last. The system allows me to do this, and I enjoy it. There is a button to show the correct answers, which I like to press last so that I can deal with it first.’ (P6)

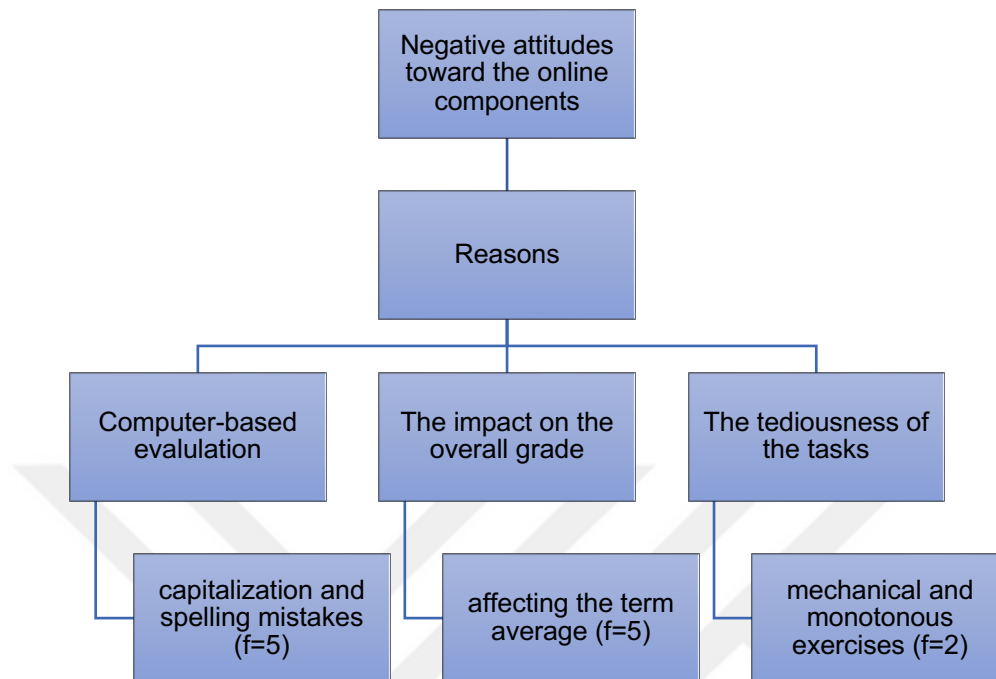
“We get feedback and may see the correct answers to the exercises that we failed. It also allows you to redo it to find the correct answer or try and fix it if you make a mistake.” (P3)

Finally, the participants included that completing the tasks in the online component is pleasant because they can be done from the comfort of their own homes. Furthermore, two of the participants indicated that even if they make technical mistakes such as capitalization or spelling mistakes, the system will accept their replies incorrectly. Interestingly, they believe it is something positive and valuable since they become more familiar with the actual exam evaluation. They both claimed that such inaccuracies are not accepted in exams in the department and that they will lose points as a result.

While two of the participants suggested that completing the exercises in the online components was time-consuming, the remaining cited some unpleasant or undesirable features that appear in the online components. The frequencies for this category are shown in Figure 19.

Figure 19

The Reasons of Negative Attitudes Toward the Online Components



Half of the participants mentioned that mechanical mistakes, such as capitalization or typos, are regarded as a valid reason not to count an answer correct, as shown in Figure 19. They consider it a disadvantage of computer-based evaluation as well as "annoying" (f=4). Furthermore, students are concerned about losing points if the system accepts some of the answers incorrectly due to the aforementioned causes, because the grade they obtain for completing the assigned tasks in the online component impacts their term average. Finally, two of the participants were critical of the online components, believing that the exercises were too mechanical and monotonous. Some of the comments are:

"I'm worried that if I don't complete some of the online exercises correctly, my score will decrease." (P5)

"The fact that the score we receive for the online component activities is factored into our term grade point average makes me a little stressed". P10

“For example, because I did not begin with a capital letter, the system recognizes what I did as incorrect, I cannot receive points, and I must edit it again, wasting my time.” (P1)

“It counts my response as incorrect if I misspell even one letter. For example, I once answered all of the questions correctly on an exercise, but the system marked it incorrect due to a spelling error. This doesn't bother me, but it is annoying because correcting it takes time.” (P8)

To conclude, according to the qualitative data, participants have a slightly positive opinion toward the online components and a noticeably negative view regarding them. The majority of them are positive about the online components because there are a variety of exercises to practice with and the opportunity for feedback for improving their current performance and enhancing their future performance. After all, it fills in the gaps where teacher feedback is not available. Cavalcanti et al. (2021) investigated the impact of automatic feedback in online learning environments and found that it improves student performance in activities. He also concluded that feedback informs student learning by displaying mistakes and successes, it causes a change in students' behavior and improves their performance, i.e., the impact of automatically provided feedback on students' performance was positive. The current study also suggests that the most notable reason that students have positive attitudes toward the online component in the current study is computerized feedback supplied by the system when students are completing tasks in the online component.

On the other hand, some participants are unwilling to complete the tasks in the online component because they worry that minor errors – as they refer to them – can be taken for incorrect replies, and correcting or redoing the exercise, as a result, is unpleasant for them. As a result, they hold a negative opinion of them. The majority of the participants do not even discuss solutions for changing negative attitudes, however, two of them proposed that the online component assignments be formatted as multiple choice and that completing the

tasks in the online components has no bearing on the term average. Falc (2013) investigated students' attitudes toward an e-textbook and found that, while many of them were frustrated by mistake clicking, the majority of them found it beneficial and appreciated the possibility to have supplementary, supporting materials such as online chapter quizzes and video libraries. The current study also found similar results as mentioned above. He also discovered that some students were easily distracted by online activities, which is in keeping with the findings of this study since some participants regarded the online component's activities to be uninteresting, monotonous, and time-consuming.



Chapter 5

Conclusion, Pedagogical Implications and Suggestions

This chapter contains the study's conclusions, pedagogical implications, and recommendations for future research.

Conclusion

It's difficult to ascertain whether students have positive or negative attitudes toward assessment and evaluation instruments because attitude has always been a complex issue. As a practical matter, positive attitudes are revealed through displays of enjoyment, joy, or pleasure, whereas negative attitudes are exposed through expressions of anxiety, worry, or apprehension in this study.

The first research question investigates the attitudes of A1-level students toward assessment and evaluation instruments. Rather than having a mostly one-sided attitude toward testing instruments, the current study's statistical findings revealed that students had both moderate positive and negative attitudes about them, with comparable mean scores. Despite the close proximity of the means, the average of the enjoyment section is only marginally higher than the anxiety section.

In the quantitative part of the study, the second and third questions investigated which of the assessment and evaluation instruments produces the most and least anxiety and enjoyment. While the speaking assessment and exams caused the most anxiety among participants, the writing portfolio and the online component caused the least anxiety. Accordingly, the highest levels of enjoyment are found in the writing portfolio and speaking assessment sections of the questionnaire, while the lowest levels of enjoyment are shown in online components and exams. The fact that the speaking exam causes both anxiety and enjoyment is noteworthy, showing that students have both negative and positive feelings about the exam.

In the qualitative part of the study, research questions 6–9 investigated what factors emerge from the assessment and evaluation instruments that influence participants' attitudes in the context of anxiety and enjoyment.

To begin, nearly all of the participants expressed both negative and positive feelings and ideas about the speaking assessment, as well as their explanations, which is in line with the quantitative findings. While linguistic challenges and paired tasks are the most commonly reported causes of negative attitudes, a sense of accomplishment is the most widely recognized cause of positive attitudes. Participants said they prefer to take the exam alone or with a partner who has a similar English level to them when it comes to tips for feeling more comfortable in speaking exams. The fact that the class instructor is also the examiner for the speaking exam is also reassuring for the students.

Second, the qualitative findings of the research question about exams revealed that participants have a favorable opinion toward quizzes but a negative attitude toward mid-term exams. This finding is consistent with the findings of the quantitative part of the study, which revealed that exams are one of the most anxiety-inducing devices for students. While unfavorable attitudes toward mid-term tests are due to the weighting of mid-term examinations in students' overall grades and the greater scope, positive attitudes regarding quizzes are due to the lesser impact of a quiz's result on the entire grade and the shorter test scope. As a moderating force for their anxiousness, the majority of participants favored the multiple-choice style for exams.

Third, the qualitative findings of the research question about the writing portfolio demonstrated that participants have a very positive attitude toward it. This is in agreement with the findings of the quantitative part of the study, which revealed that the writing portfolio is one of the highest means from the enjoyment section of the questionnaire. While receiving feedback from the teacher is one of the most noticeable reasons for this positive attitude because it is regarded as a valuable component that can guide and provide an opportunity to gradually improve one's writing skills, writing under time constraints and controlled tasks

are the explanations for negative attitudes toward the writing portfolio. Participants cite increased time allocation and more flexible tasks as elements that can enhance enjoyment.

Lastly, the qualitative findings of the research question about the online component demonstrated that the online components receive a slightly favorable response from participants. In other words, the online components have a slightly positive and noticeably negative image among participants. This finding is consistent with the quantitative findings, which revealed that the online component has one of the lowest mean scores in the questionnaire's enjoyment section. The participants have positive views since they have a variety of activities to practice with and the ability to get feedback on how to improve their present performance. Computer-based evaluation, on the other hand, which shows a highly rigorous evaluation of correct and incorrect answers, as well as the impact of the grade obtained from the online components, are both common sources of negative attitudes. It's worth noting that students prefer adjectives like "time-wasting, annoying, and not enjoyable" to describe their feelings and ideas about this instrument, implying that it causes dissatisfaction and boredom rather than anxiety among participants.

The fourth and fifth research questions investigated if there is a significant relationship between gender and academic major in the attitudes of A1-level students toward testing instruments in terms of the extent of anxiety and enjoyment. There are two sections to the questionnaire: anxiety and enjoyment. In the anxiety section, the results demonstrate that gender makes a significant difference in attitudes toward assessment and evaluation tools, with female students being more nervous than male students. This demonstrates that female students are more apprehensive, indicating a negative attitude toward testing tools. However, there was no significant difference between male and female students in the enjoyment section of the questionnaire, indicating that gender differences had no effect on participants' enjoyment levels. In both sections of the questionnaire, however, the results revealed that an academic major makes a significant difference in attitudes about assessment and evaluation instruments. When compared to students

majoring in Natural Sciences, students majoring in Social Sciences appear to be more anxious and enjoy less when taking tests (i.e. quizzes, midterm exams, and speaking exams) and completing activities in the writing portfolio and online component, indicating that while students majoring in Social Sciences have negative attitudes toward assessment and evaluation instruments, students majoring in Natural Sciences have more positive attitudes toward assessment and evaluation instruments.

When all of the research data are taken into account, each student has their interpretations, perceptions, feelings, and beliefs about each assessment instrument. Former experiences, context, personal interest in specific language abilities, and the components of each instrument all contribute to the student's unique approach to the testing instruments, which can alter regularly. Students' attitudes about assessment and their value for educational procedures should be viewed with caution in this manner. The majority of study findings, on the other hand, reveal patterns, tendencies, and some shared concerns between students' attitudes and various assessment and evaluation instruments (Struyven, et al., 2005). As a result, the data gathered from students can provide useful insights for instructors, testers, and students themselves, though the data is to be enhanced by additional supporting components such as teachers' and testers' perspectives, attitudes of students at various levels of English, and the students' academic achievement scores.

Pedagogical Implications

We can help students feel more secure and less afraid by recognizing, understanding, and responding to the underlying processes that underpin their anxious attitudes and behaviors (Bledsoe & Baskin, 2014). The results of this study show that there are various reasons causing negative attitudes and feelings about testing instruments as well as positive ones. While it is critical to understand the sources of positive attitudes in order to improve and sustain them, it is also essential to identify and address the causes of negative attitudes. Negative attitudes pose a threat to students' enhanced performance in

this regard. As Bledsoe and Baskin (2014) argued that continued fear-based emotions would most certainly overload the central nervous system, resulting in a cascade of unnecessary stress and its unfavorable consequences for learning. Thus, educators can employ the following strategies:

- *Learning about anxiety and its effects on students:* Because negative emotions like anxiety and fear can manifest in a variety of ways, it's not always easy to spot them, especially during exams or assessments, which are commonly regarded as potentially anxiety-inducing situations. As a result, research aimed at revealing students' attitudes and the causes behind those attitudes might serve as a foundation and raise awareness.
- *Teaching students about the anxiety that they might experience during tests or evaluations, as well as how to cope:* Students must notice and be aware of their attitudes in order to deal with negative attitudes. If a student has negative self-talk before or during a test, the instructor can encourage him/her to investigate the reasons for his/her behavior, identify problem areas, and seek help (e.g., recognize gaps, ask for office hours, consult student counselor). Instead of feeling anxious and helpless, the learner will be able to produce answers that will aid him/her.
- *Creating a supportive environment for students:* It's vital to address students' requirements, learn about their opinions, and seek feedback on the testing procedures from them when evaluating their performance. Students will notice that the department uses more individualized and learner-centered assessment procedures, which can help to create a more pleasant and anxiety-free environment.

In addition to what is recommended above, knowing students' attitudes has enabled us to better construct programs focused on meeting student needs and assessing their English knowledge because existing testing instruments are meant to check and evaluate whether the objectives are met and students truly learn. Examining student attitudes is also efficient to reveal the assessment and evaluation instruments used to assess students'

performance throughout the academic term to enhance students' educational experiences and learning. This information has also aided us in making improvements to our program and, as a result, to our testing devices. In other words, in our drive to understand student learning, the students' thoughts and feelings of assessment are valuable feedback as student attitudes assist us in our reflective efforts to enhance educational practices and provide a greater level of learning and education for our students.

In this study, students expressed how they felt about testing instruments by labeling their feelings as anxious, tense, or stressed, which helps to disclose their negative attitudes toward assessment and evaluation instruments. According to Burklund et al (2014), affect labeling in everyday life includes naming one's own affective reactions to specific events or items, as well as labeling emotions as angry, worried, nervous, or sad, even if it doesn't involve an active intention of modifying felt emotions. They discussed it being a type of unintended emotion-focused coping that leads to less anger, anxiety, and worry.

While curiosity and amusement engage attention and foster greater involvement, boredom, anxiety, and fear interrupt concentration and interfere with learning (Brackett & Simmons, 2015). Students' feelings and attitudes can affect their performance negatively or positively in testing, just as they can in learning. That's why testing units can use how students feel and think about different assessment instruments used at the department to make changes if necessary, such as replacing some components or changing the task or test specifications to pave the way for students to perform better. Testing units may also need to keep in mind the domain chosen for the general language learning and teaching goals indicated in the CEFR when choosing scenarios, purposes, tasks, themes, and texts for teaching and testing materials and activities (Council of Europe, 2001).

This research can serve as a starting point for educators and administrators. If educators can initiate discussions with students who are experiencing high levels of anxiety as a result of any assessment or evaluation instrument, they can attempt to reduce tensions before they become persistent and cause academic and emotional stress.

It's also worth noting that this research focuses on students' positive attitudes and the reasons behind their positive thoughts. There is reason to believe that studying positive emotions in more depth will give fresh insights into the systems involved in language, as Dewaele and MacIntyre (2014) pointed out, and interest in positive emotions has surged in recent years. Students have both negative and good feelings and attitudes regarding testing instruments, and the reasons for favorable attitudes are addressed and provided in-depth in the qualitative section. In this regard, it appears necessary to emphasize those pleasant emotions can help to mitigate the effects of negative emotional arousal (Dewaele & MacIntyre, 2014).

As a result of this research, testers can gain insight into which instruments induce anxiety and enjoyment, which can be a very useful tool when setting up the testing instrument. For example, if testers know how difficult a test/task is for students and whether the test or task's challenge level surpasses students' abilities, generating anxiety or fear, they can use this information to encourage enjoyment and flow experiences by creating the appropriate arrangements for the testing instrument. It's, therefore, crucial to know whether students are enjoying the testing instruments and their various components or whether they are anxious.

Alderson and Wall (1993) stated that exams influence teaching, learning, how and what teachers teach, and how and what students learn, according to the washback hypothesis (as cited in Bailey, 1994, p.263). If departments gain knowledge of how students feel about testing instruments, which may also provide conscious feedback, they may be inclined to make changes to the assessment instruments in terms of test and task criteria, skills included, and percentages for assessment components. According to Morrow (1991), when this is done on a regular basis during different time periods, such as the first and second terms, a conscious feedback loop between testing and teaching will be achievable, which is a critical mechanism for educational progress (as cited in Bailey, 1994, p.261). It can also be put forward that this conscious feedback loop could also provide valuable and

meaningful insights for testing units in basic English departments when deciding on term percentages and assessment instruments.

To conclude, even though the instrument choices and requirements have already been set by the educators, students have the opportunity to voice their feelings about how they are tested, suggest alternatives, and express their opinions concerning testing instruments. This makes it feasible to see whether students favor or dislike the use of implemented testing instruments.

Suggestions

- The achievement of students was not included in this study. In addition to their attitudes, their scores and performance on assessment and evaluation instruments can provide information on the testing instruments' washback effects.
- Actual classroom observation can be integrated to better discover students' attitudes because physical reactions to tests and evaluations can be observed.
- The age of the students was not taken into account in this study because the age range in university preparatory programs does not exhibit considerable variation. As a result, a comparable study might be undertaken with students of various ages to examine whether age has an impact on students' views toward assessment and evaluation tools in terms of anxiety and enjoyment.
- Students who are A1-level university prep classes are the subjects of the current study; however, other English proficiency levels could be included in a future study because students with higher English proficiency levels may show different levels of anxiety and enjoyment.
- Since this study was completed in the fall semester, another study might be undertaken in the spring and fall semesters to evaluate students' opinions toward testing instruments. This recommendation is based on the fact that assessment and

evaluation instruments, as well as their components such as task and test types, can differ in academic terms. Concerning this suggestion, it should be noted that this study was conducted in the fall semester of 2021-2022 when the lessons were delivered via distance education, and as a result, the testing instruments did not assess listening skills; therefore, another study can be conducted when all other four skills have been assessed. Furthermore, students' attitudes toward testing instruments may differ depending on whether they are taking online or face-to-face tests.



References

- Akdamar, H. S. (2020). *Students' attitudes towards English in the Turkish secondary school context*. (Publication No: 636063) [Master's Thesis, Çağ University].YOK National Thesis Center.
- Aida, Y. (1994). Examination of Horwitz, Horwitz, and Cope's construct of foreign language anxiety: The case of students of Japanese. *The Modern Language Journal*, 78(2) 155-167. <https://doi.org/10.2307/329005>
- Alderson, J. C. & Wall, D. (1993). *Does washback exist?*. *Applied Linguistics*, 14(2), 115–129. Doi:10.1093/applin/14.2.115.
- Allen, J. D. (2005). Grades as valid measures of academic achievement of classroom learning. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 78(5), 218-223. Doi: 10.3200/TCHS.78.5.218-223.
- Arabai, F. (2014). The influence of teachers' anxiety-reducing strategies on learners' foreign language anxiety. *Innovation in Language Learning and Teaching*, 1-28. Doi: 10.1080/17501229.2014.890203.
- Altay, İ. F. (2007). Pragmatics of testing. *Journal of Language and Linguistic Studies*, 3(2),266-288. License: [CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/)
- Altunışık, R. (2008). Anketlerde veri kalitesinin iyileştirilmesi için öntest (pilot) yöntemleri. *Pazarlama ve Pazarlama Araştırmaları Dergisi*, 1(2),1-17.
- Aragão, R. (2011). Beliefs and emotions in foreign language learning. *System* 39(3),302–313. Doi:10.1016/j.system.2011.07.003.
- Armstrong, M. C. F., Marin, M. I. R., & Patarroyo, L. M. (2004). Analysis of assessment instruments used in foreign language teaching. *Ikala, revista de Lenguaje y Cultura*, 9 (15), 170-201.

- Arnold, J. (2011). Attention to affect in language learning. *Anglistik, International Journal of English Studies*, 22(1), 11-22. ERIC Number: ED532410.
- Aryadoust, V., Eckes, T., & In'nami Y. (2021) Editorial: Frontiers in language assessment and testing. *Frontiers in Psychology*, 12(69), 1-4. Doi:10.3389/fpsyg.2021.691614.
- Aydın, S. (2008). An investigation on the language anxiety and fear of negative evaluation among Turkish EFL learners. *Asian EFL Journal*, 421-44.
- Azarfam, A. Y., Samad, A. A., & Noordin, N. (2016). Exploring EFL learners' attitudes toward the application of a model of writing e-portfolio. *International Journal of Applied Linguistics and English Literature*, 5(3), 57-68. Doi: [10.7575/aiac.ijalel.v.5n.3p.57](https://doi.org/10.7575/aiac.ijalel.v.5n.3p.57)
- Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, 13(3), 257-279. Doi:10.1177/026553229601300303.
- Bata, S., & Castro, C. (2021). English as a foreign language students' emotional intelligence management when taking speaking exams. *Profile Issues in Teachers Professional Development*, 23(2), 245-261. <https://doi.org/10.15446/profile.v23n2.88378>.
- Baltacı, A. (2018). Nitel araştırmalarda örnekleme yöntemleri ve örnek hacmi sorunsalı üzerine kavramsal bir inceleme. *Bitlis Eren Üniversitesi Sosyal Bilimler Dergisi*, 7(1), 231-274.
- Berlak, H. (1992). The need for a new science of assessment. In H. Berlak, F. M. Newman, E. Adams, D. A. Archbald, T. Burgess, J. & Raven, T. A. Romberg (Eds.), *Toward a new science of educational testing and assessment* (pp.1-22). State University of New York Press.
- Blaikie, F., Schönau, D., & Steers, J. (2004). Preparing for portfolio assessment in art and design: A study of the opinions and experiences of exiting secondary school students in Canada, England and The Netherlands. *International Journal of Art & Design Education*, 23(3), 302-315.

- Bledsoe, T. S., & Baskin, J. J. (2014). Recognizing student fear: The elephant in the classroom. *College Teaching*, 62(1), 32-41. <https://doi.org/10.1080/87567555.2013.831022>
- Botes, E., Dewaele, J. M., & Greiff, S. (2021a). Taking stock: An overview of the literature and a preliminary meta-analysis of Foreign Language Enjoyment and other individual difference variables. <https://doi.org/10.31234/osf.io/quaj5>.
- Botes, E., Dewaele, J.-M., & Greiff, S. (2022). Taking stock: A meta-analysis of the effects of foreign language enjoyment. *Studies in Second Language Learning and Teaching*, 12(2), 205-232. <https://doi.org/10.14746/ssllt.2022.12.2.3>
- Brackett, M. A., & Simmons, D. (2015). Emotions matter. *Educational leadership*, 73(2), 22-27.
- Brown, H. D. (2003). *Principles of language learning and teaching*. Longman.
- Brown, J. D. & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly*, 32(4), 653-675. <https://doi.org/10.2307/3587999>.
- Burgess, S. & Head, S. (2005). *How to teach for exams* (J. Harmer, Ed.). Pearson Education Limited.
- Burklund, L. J., Creswell, J. D., Irwin, M., & Lieberman, M. (2014). The common and distinct neural bases of affect labeling and reappraisal in healthy adults. *Frontiers in Psychology*, 5, 221. <https://doi.org/10.3389/fpsyg.2014.00221>
- Carr, N. T. (2011). *Designing and analyzing language tests*. Oxford University Press.
- Cavalcanti, A. P., Barbosa, A., Carvalho, R., Freitas, F., Tsai, Y. S., Gašević, D., & Mello, R. F. (2021). Automatic feedback in online learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence*, 2, 100027, 1-17. <https://doi.org/10.1016/j.caeai.2021.100027>

- Chang, G. B. (1996). A study of anxiety in Chinese EFL learners. *Teaching and Research*, 18, 68-90.
- Cheng, Y. S. (2002). Factors associated with foreign language writing anxiety. *Foreign Language Annals*, 35(6), 647-656. Doi: 10.1111/j.1944-9720.2002.tb01903.x.
- Cheng, Y. S., Horwitz, E. K., & Schallert, D. L. (1999). Language anxiety: Differentiating writing and speaking components. *Language Learning*, 49(3), 417-446. Doi: 10.1111/0023-8333.00095.
- Čiček, L. (2015). *Language learning anxiety-causes and consequences*. [Unpublished doctoral dissertation] University of Zagreb.
- Cohen, A. D. (1984). On taking language tests. *Language Testing*, 1(1), 70–81. Doi:10.1177/026553228400100106.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment (CEFR)*. Cambridge, U.K: Press Syndicate of the University of Cambridge.
- Creswell, J. W. (2009). *Research design* (3rd ed.). SAGE Publications.
- Crowe, S., Cresswell, K., Robertson, A., Hubby, G., Avery, A., & Sheikh, A., (2011). The case study approach. *BMC Medical Research Methodology*, 11(1).1-9. Doi: 10.1186/1471-2288-11-100.
- Cumming, A. (2009). Language assessment in education: tests, curricula, and teaching. *Annual Review of Applied Linguistics*, 29, 90–100. doi:10.1017/s0267190509090084.
- Dalkılıç, N. (2001). *An investigation into the role of anxiety in second language learning*. [Unpublished doctoral dissertation] Çukurova University.

- Davidson, S. (2018). *A Multi-dimensional model of enjoyment: Development and validation of an enjoyment scale (Enjoy)*. [Unpublished doctoral dissertation] Embry-Riddle Aeronautical University.
- Demirdağ, Ö. (2012). *Foreign language anxiety and performance of language students in Turkish University preparatory classes*. (Publication No: 319987) [Master's Thesis, Bolu Abant İzzet Baysal University].YOK National Thesis Center.
- Deretarası, Ö. (2020). *Investigating EFL instructors' language assessment knowledge*. (Publication No: 648977) [Master's Thesis, Dokuz Eylül University].YOK National Thesis Center.
- Dewaele, J. M., & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning and Teaching*, 4(2), 237-274. Doi: 10.14746/ssllt.2014.4.2.5.
- Dewaele, J. M., Özdemir, C., Karci, D., Uysal, S., Özdemir, E. D., & Balta, N. (2019). How distinctive is the foreign language enjoyment and foreign language classroom anxiety of Kazakh learners of Turkish?. *Applied Linguistics Review*, 1-23. <https://doi.org/10.1515/applirev-2019-0021>
- Dewaele, J. M., Witney, J., Saito, K., & Dewaele, L. (2018). Foreign language enjoyment and anxiety: The effect of teacher and learner variables. *Language Teaching Research*, 22(6), 676-697. Doi: .oOr.g1/107.171/1773/61236126186818717669922161
- Doyle, A. (2019). What is a semi-structured interview. Retrieved from <https://www.thebalancecareers.com/what-is-a-semi-structured-interview-2061632>.
- Falc, E. O. (2013). An assessment of college students' attitudes towards using an online e-textbook. *Interdisciplinary Journal of E-Learning and Learning Objects*, 9(1), 1-12.
- Frodden Armstrong, M. C., Restrepo Marín, M. I. & Maturana Patarroyo, L. (2004). Analysis of assessment instruments used in foreign language teaching. *Íkala*. 9(1), 171 – 201.

https://www.researchgate.net/publication/277872114_Analysis_of_Assessment_Instruments_Used_in_Foreign_Language_Teaching.

- Gökhan, Ç. (2004). *Attitudes of teachers and students towards the assessment system at Hacettepe University Basic English Department and their opinions about alternative methods of assessment*. (Publication No: 145551) [Master's Thesis, Bilkent University].YOK National Thesis Center.
- Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold.
- Gultom, E. (2016). Assessment and evaluation in EFL teaching and learning. *Proceedings of the Fourth International Seminar on English Language and Teaching (ISELT-4)*, 190-198. ISBN: 978-602-74437-0-9.
- Harmer, J. (2004). *How to teach writing*. Pearson Education Limited.
- Harmer, J. (2007). *How to teach English*. Pearson Education Limited.
- Horwitz, E. K. (1986). Preliminary evidence for the reliability and validity of a foreign language anxiety scale. *TESOL Quarterly*, 20,(3), 559-562.
<http://www.jstor.org/stable/3586302>
- Horwitz, E. K., Horwitz, M. B., & Cope, J. A. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125–132. <https://doi.org/10.2307/327317>.
- Horwitz, E. K. (2010). Foreign and second language anxiety. *Language Teaching*, 43(2), 154 – 167. DOI: <https://doi.org/10.1017/S026144480999036X>
- Hughes, A. (1989). *Testing for language teachers*. Cambridge University Press.
- Huynh, T. (2021). A complex dynamic systems approach to foreign language learners' anxiety in the emergency online language classrooms. *Computer Assisted Language Learning Electronic Journal (CALL-EJ)*, 22(3), 200-229.
<https://www.researchgate.net/publication/354065986>.

- Kayhan, E. (2021). *Young learners' foreign language anxiety level and anxiety sources*. [Unpublished master's thesis] Sakarya University.
- Kráľová, Z. (2016). *Foreign language anxiety*. Constantine the Philosopher University in Nitra.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.
- Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement*, 30(3), 607-610. <https://doi.org/10.1177/001316447003000308>
- Leo, D. & Sardanelli, F. G. (2020) Statistical significance: *p* value, 0.05 threshold, and applications to radiomics—reasons for a conservative approach. *European Radiology Experimental*, 18(4). <https://doi.org/10.1186/s41747-020-0145-y>.
- Li, C. (2020). A positive psychology perspective on Chinese EFL students' trait emotional intelligence, foreign language enjoyment and EFL learning achievement. *Journal of Multilingual and Multicultural Development*, 41(3), 246-263. Doi: 10.1080/01434632.2019.1614187.
- Lynch, B. K. (2001). Rethinking assessment from a critical perspective. *Language Testing*, 18(4), 351-372. <https://doi.org/10.1177/026553220101800403>.
- MacIntyre, P. D. (2017). An overview of language anxiety research and trends in its development. In C. Gkonou, M. Daubney, & J.-M. Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications* (pp. 11–30). Multilingual Matters.
- MacIntyre, P. D., & Gardner, R. C. (1991). Language anxiety: Its relationship to other anxieties and to processing in native and second languages. *Language Learning*, 41, 513-534. https://petermacintyre.weebly.com/uploads/1/0/1/8/10187707/language_anxiety1991.pdf

- MacIntyre, P. D., & Mercer, S. (2014). Introducing positive psychology to SLA. *Studies in Second Language Learning and Teaching*, 4(2), 153-172. Doi: 10.14746/ssllt.2014.4.2.2.
- Marshall, M. N. (1996). Sampling for qualitative research. *Family Practice*, 3(6), 522-525.
- McDonough, K., & Fuentes, C. G. (2015). The effect of writing task and task conditions on Colombian EFL learners' language use. *TESL Canada Journal*, 32(2), 67-67. <https://doi.org/10.18806/tesl.v32i2.1208>.
- Meidasari, V. E. (2015). The assessment and evaluation in teaching English as a foreign language. *Indonesian EFL Journal*, 1(2), 224-231.
- Mierzwa, E. (2018). The relationship between foreign language enjoyment and gender among secondary grammar school students. *The Journal of Education, Culture, and Society*, (2), 117-135. Doi: 10.15503/jecs20182.117.135
- Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis* (2nd ed.). SAGE publications.
- Moore, L. L. (2010). Students' attitudes and perceptions about the use of cooperative exams in an introductory leadership class. *Journal of Leadership Education*, 9(2), 72-85.
- Nath, P.R., Mohamad, M., & Yamat, H. (2017). The effects of movies on the affective filter and English acquisition of low-achieving English learners. [Abstract]. *Creative Education*, 8(8).
- Onwuegbuzie, A. J., Bailey, P., & Daley, C. E. (1999). Factors associated with foreign language anxiety. *Applied Psycholinguistics*, 20(2), 217–239. <https://doi.org/10.1017/S0142716499002039>
- Orcan, F. (2020). Parametric or non-parametric: Skewness to test normality for mean comparison. *International Journal of Assessment Tools in Education*, 7(2), 255-265. <https://doi.org/10.21449/ijate.656077>

- Overton, T. (2012). *Assessing learners with special needs* (7th ed.). Pearson.
- Özer, Z. & Altay, İ. F. (2021). Examining the level of enjoyment and anxiety among Turkish EFL students. *Journal of Language and Linguistic Studies*, 17(Special Issue 1), 663-671. ISSN: 1305-578X.
- Paker, T., & Höl, D. (2012). Attitudes and perceptions of the students and instructors towards testing speaking communicatively. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 32,13-24. <http://hdl.handle.net/11499/3750>.
- Pappamihel, N.E. (2002). English as a second language students and English language anxiety: Issues in the mainstream classroom. *Research in the Teaching of English*, 36(3), 327-355. http://lchc.ucsd.edu/mca/Mail/xmcamail.2004_09.dir/att-0070/01-RT0363English.pdf.
- Powers, D. E. (2010). The case for a comprehensive, four-skills assessment of English-language proficiency. *R&D Connections*, 14, 1-12. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.599.5802&rep=rep1&type=pdf>
- Rao, P. S. (2018). Evaluation and assessment in ELT: A global perspective. *Research Journal of English Language and Literature (RJELAL)*, 6(2), 337-346. Doi:10.33329/rjelal.62.237.
- Reid, N. (2015). Attitude research in science education. In M. S. Khine (Ed.), *Attitude measurements in science education: Classic and contemporary approaches* (pp.3-46). Information Age Publishing.
- Reynolds, N., Diamantopoulos, A., & Schlegelmilch, B. (1993). Pre-testing in questionnaire design: a review of the literature and suggestions for further research. *Market Research Society. Journal.*, 35(2), 1-11.

- Sarason, I. G. (1984). Stress, anxiety, and cognitive interference: Reactions to tests. *Journal of Personality and Social Psychology*, 46(4), 929–938. <https://doi.org/10.1037/0022-3514.46.4.929>.
- Sari, N. P., & Arsyad, M. (2021). Environment and differences of self-adjustment ability between students of natural sciences programs and students of social sciences programs. *Advances in Social Science, Education and Humanities Research*, 525, 315-321. <http://creativecommons.org/licenses/by-nc/4.0/>. 3.
- Scovel, T. (1978). The effect of affect on foreign language learning: A review of the anxiety research. *Language Learning*, 28(1), 129–142. Doi:10.1111/j.1467-1770.1978.tb00309.x
- Segers, M. & Dochy, F. (2001) New assessment forms in problem-based learning: the value-added of the students' perspective, *Studies in Higher Education*, 26(3), 327–343. <https://doi.org/10.1080/03075070120076291>.
- Shohamy, E. (1993). *The power of tests: The impact of language tests on teaching and learning*. NFLC Occasional Papers. Johns Hopkins Univ., Washington, DC. National Foreign Language Center. ERIC Number: ED362040.
- Slater, T. F. (1996). Portfolio assessment strategies for grading first-year university physics students in the USA. *Physics Education*, 31(5), 329.
- Spolsky, B. (2016). History of language testing. *Language Testing and Assessment*, 1–11. Doi:10.1007/978-3-319-02326-7_32-1.
- Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: A review. *Assessment & Evaluation in Higher Education*, 30(4), 325-341. <https://doi.org/10.1080/02602930500099102>.
- Swamy, S. (2007). How should you frame questions to measure user attitudes accurately? An experimental design study. In N. Aykin (Ed.), *Usability and internationalization* (pp.496-505). Global and Local User Interfaces. UI-HCII 2007. Lecture Notes in

Computer Science, 4560. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-73289-1_57.

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in science education*, 48(6), 1273-1296. Doi: 10.1007/s11165-016-9602-2.

Tamunobelem, I. (2016). Testing as an inevitable instrument in today's language learning environment. *AFRREV IJAH: An International Journal of Arts and Humanities*, 5(2), 216-224. Doi: 10.4314/ijah.v5i2.17.

Tamunobelem, I. (n.d). The roles of a 21st century language teacher in the classroom. Academia. https://www.academia.edu/31697122/Testing_as_an_inevitable_instrument_in_todays_language_learning_environment.

Tezbaşaran, A. A. (2008). Likert tipi ölçek hazırlama kılavuzu. (Üçüncü Sürüm) Türk Psikologlar Derneği. Retrieved from https://www.academia.edu/1288035/Likert_Tipi_%C3%96l%C3%A7ek_Haz%C4%B1rlama_K%C4%B1lavuzu

Thomas, S. D., Hathaway, D. K., & Arheart, K. L. (1992). Face validity. *Western Journal of Nursing Research*, 14(1), 109–112. Doi:10.1177/019394599201400111.

Tosuncuoğlu, İ. (2018). Importance of assessment in ELT. *Journal of Education and Training Studies*, 6(9), 163-167. Doi: 10.11114/jets.v6i9.3443.

Trifoni, A., & Shahini, M. (2011). How does exam anxiety affect the performance of university students?. *Mediterranean Journal of Social Sciences*, 2(2), 93-93. ISSN online: 2039-2117.

Tuncer, M. & Doğan, Y. (2015). Effect of foreign language classroom anxiety on Turkish university students' academic achievement in foreign language learning. *Journal of Education and Training Studies*, 3(6), 14-19. Doi: 10.11114/jets.v3i6.998.

- Wörde, R. (2003). Students' perspectives on foreign language anxiety. *Inquiry*, 8(1), 1-15.
<http://www.vccaedu.org/inquiry/inquiry-spring2003/i-81-worde.html>
- Young, D. J. (1991). Creating a low-anxiety classroom environment: What does the anxiety research suggest?. *Modern Language Journal*, 75, 426-439. Doi:10.2307/329492.
- Youssef, A. M. S. (2012). Role of motivation and attitude in introduction and learning of English as a foreign language in Libyan high schools. *International Journal of Linguistics*, 4(2), 366-375. Doi:10.5296/ijl.v4i2.1855.
- Yurdugül, H. & Bayrak, F. (2012). Content validity measures in scale development studies: Comparison of content validity index and kappa statics.[Abstract] *Hacettepe University Journal of Education*, 264-271.
- Yüce, Z. (2015). *Pre-service English language teachers' conceptions of assessment and assessment practices*. (Publication No: 379999) [Master's Thesis, Çağ University].YOK National Thesis Center.
- Ziyan Atlı, Z. (2017). *Foreign language anxiety in English language learning*. [Unpublished master's thesis] İstanbul Aydın University.

APPENDIX-A: METU Human Subjects Ethics Committee Institutional Approval Form**İnsan Araştırmaları Etik Merkezi**

RE: Etik kurul izni başvurusu

Kime: İnsan Araştırmaları Etik Merkezi, Ayca İrem Yılmaz,

Bilgi: İsmail Fırat Altay

15:14

[Gizle](#)**Sayın Dr.Öğretim Üyesi İsmail Fırat ALTAY**

Danışmanlığını yürüttüğünüz Dr.Öğr.Gör. Ayca İrem YILMAZ'ın “Üniversite İngilizce Hazırlık Sınıfı Öğrencilerinin Ölçme ve Değerlendirme Araçlarına Yönelik Tutumları” başlıklı araştırması Hacettepe Üniversitesinden aldığınız Etik Onay doğrultusunda üniversitemizde araştırma yapmanız uygun bulunmuştur.

Bilgilerinize**Saygılarımla***Aysel KÖSEGİL**İnsan Araştırmaları**Etik Kurulu Koordinatörlüğü (İAEK)*

APPENDIX-B: The Questionnaire

Questionnaire Items					
	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1. Speaking exams make me nervous.					
2. I feel stressed over not making proper sentences and explaining my ideas in speaking exams.					
3. I feel nervous in speaking exams as I feel anxious over making a mistake in front of the instructor.					
4. I feel stressed in speaking exams if the instructor isn't relaxed.					
5. I feel nervous in speaking exams when we perform the task in pairs.					
6. I feel stressed in speaking exams when I perform the task individually.					
7. Quizzes make me nervous.					
8. Midterm exams make me nervous.					
9. I feel anxious in quizzes and midterm exams as I experience anxiety-provoking self-talk. (e.g. I can't do it. I will forget everything etc.)					
10. I feel stressed in quizzes and midterm exams because of the time limit.					
11. I feel stressed in quizzes and mid-term exams if the instructor isn't relaxed.					
12. Completing the tasks in the writing portfolio makes me nervous.					
13. I feel nervous while completing the tasks in the writing portfolio as I have some linguistic problems (inadequate mastery of vocabulary and sentence structures, grammatical errors etc.)					
14. I feel nervous about the negative comments and evaluation of the instructor about my work in the writing portfolio.					
15. I feel stressed when I have to write under time constraints.					
16. Doing the exercises in the online component (online platform of the coursebook) makes me nervous.					

17. I feel anxious about doing the exercises in the online component as it is artificial intelligence that evaluates my work.					
18. I feel stressed about completing the tasks in the online component as there is a deadline.					
19. I enjoy speaking exams.					
20. I take pleasure in the speaking exam when the topic of the task is relevant to our interests.					
21. I enjoy the speaking exam when we perform the task in pairs.					
22. I enjoy the speaking exam when I perform the task individually.					
23. I take pleasure in the speaking exam when I'm praised for my good performance.					
24. I enjoy the speaking exam if the instructor's attitude is positive.					
25. I enjoy quizzes.					
26. I enjoy mid-term exams.					
27. I take pleasure in taking the quizzes and mid-term exams since they help me to identify my gaps.					
28. I enjoy completing the tasks in writing portfolio.					
29. I enjoy completing the tasks in writing portfolio as it helps me to improve my writing skills gradually.					
30. I enjoy doing the exercises in the online component.					
31. I am pleased by completing the tasks in the online component as I can do extra exercises.					
32. I enjoy doing the exercises in online component as I can do them at my own pace.					

APPENDIX-C: Semi-Structured Interview Form**INTERVIEW QUESTIONS**

1. How do you feel about speaking exams? Can you describe your feelings?
 - a. What do you like best about it? What factors do you believe contribute to your enjoyment level?
 - b. What disturbs you the most? What factors can you name as having potential to evoke anxiety for you?
 - i. What can help you to reduce your anxiety in speaking exams? (If there are factors causing anxiety)
2. How do you feel about quizzes and mid-term exams? Can you describe your feelings?
 - a. What do you like best about it? What factors do you believe contribute to your enjoyment level?
 - b. What disturbs you the most? What factors can you name as having potential to evoke anxiety for you?
 - i. What can help you to reduce your anxiety in quizzes and mid-term exams? (If there are factors causing anxiety)
3. How do you feel about completing the tasks in the writing portfolio? Can you describe your feelings?
 - a. What do you like best about it? What factors do you believe contribute to your enjoyment level?
 - b. What disturbs you the most? What factors can you name as having potential to evoke anxiety for you?
 - i. What can help you to reduce your anxiety about completing the tasks in the writing portfolio? (If there are factors causing anxiety)
4. How do you feel about doing the exercises in the online component? Can you describe your feelings?
 - a. What do you like best about it? What factors do you believe contribute to your enjoyment level?
 - b. What disturbs you the most? What factors can you name as having potential to evoke anxiety for you?
 - i. What can help you to reduce your anxiety about doing the exercises in the online component? (If there are factors causing anxiety)

APPENDIX-D: Ethics Committee Approval



T.C.
HACETTEPE ÜNİVERSİTESİ REKTÖRLÜĞÜ
Rektörlük

Tarih: 17/01/2022
Sayı: E-35853172-300-00001978340
00001978340

Sayı : E-35853172-300-00001978340
Konu : Ayça İrem YILMAZ (Etik Komisyon İzni)

17.01.2022

EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İlgi : 17.12.2021 tarihli ve E-51944218-300-00001926141 sayılı yazınız.

Enstitünüz Yabancı Diller Eğitimi Anabilim Dalı İngiliz Dili Eğitimi Yüksek Lisans Programı öğrencilerinden **Ayça İrem YILMAZ**'ın **Dr. Öğr. Üyesi İsmail Fırat ALTAY** danışmanlığında yürüttüğü "**Üniversite İngilizce Hazırlık Sınıfı Öğrencilerinin Ölçme ve Değerlendirme Araçlarına Yönelik Tutumları**" başlıklı tez çalışması Üniversitemiz Senatosu Etik Komisyonunun **11 Ocak 2022** tarihinde yapmış olduğu toplantıda incelenmiş olup, etik açıdan uygun bulunmuştur.

Bilgilerinizi ve gereğini rica ederim.

Prof. Dr. Vural GÖKMEN
Rektör Yardımcısı

Bu belge güvenli elektronik imza ile imzalanmıştır.

Belge Doğrulama Kodu: 290E7246-8C4C-42C6-B3EE-7034684E8E5D

Belge Doğrulama Adresi: <https://www.turkiye.gov.tr/hu-ebys>

Adres: Hacettepe Üniversitesi Rektörlük 06100 Sıhhiye-Ankara

Bilgi için: Sevdâ TOPAL

E-posta: yazimd@hacettepe.edu.tr İnternet Adresi: www.hacettepe.edu.tr Elektronik

Bilgisayar İşletmeni

Ağ: www.hacettepe.edu.tr

Telefon: 0 (312) 305 3001-3002 Faks: 0 (312) 311 9992

Telefon: 03123051008

Kep: hacettepeuniversitesi@hs01.kep.tr



APPENDIX-E: Declaration of Ethical Conduct

I hereby declare that...

- I have prepared this thesis in accordance with the thesis writing guidelines of the Graduate School of Educational Sciences of Hacettepe University;
- all information and documents in the thesis/dissertation have been obtained in accordance with academic regulations;
- all audio visual and written information and results have been presented in compliance with scientific and ethical standards;
- in case of using other people's work, related studies have been cited in accordance with scientific and ethical standards;
- all cited studies have been fully and decently referenced and included in the list of References;
- I did not do any distortion and/or manipulation on the data set,
- and **NO** part of this work was presented as a part of any other thesis study at this or any other university.

18 /07/2022

Ayça İrem YILMAZ

APPENDIX-F: Thesis/Dissertation Originality Report

18/07/2022

HACETTEPE UNIVERSITY
Graduate School of Educational Sciences
To The Department of Foreign Languages Education

Thesis Title: University English Preparatory School Students' Attitudes Towards Assessment and Evaluation Instruments

The whole thesis that includes the *title page, introduction, main chapters, conclusions and bibliography section* is checked by using **Turnitin** plagiarism detection software take into the consideration requested filtering options. According to the originality report obtained data are as below.

Time Submitted	Page Count	Character Count	Date of Thesis Defense	Similarity Index	Submission ID
06/07/2022	110	169,606	17/06/2022	%12	1843874961

Filtering options applied:

1. Bibliography excluded
2. Quotes included
3. Match size up to 5 words excluded

I declare that I have carefully read Hacettepe University Graduate School of Educational Sciences Guidelines for Obtaining and Using Thesis Originality Reports; that according to the maximum similarity index values specified in the Guidelines, my thesis does not include any form of plagiarism; that in any future detection of possible infringement of the regulations I accept all legal responsibility; and that all the information I have provided is correct to the best of my knowledge.

I respectfully submit this for approval.

Name Lastname: Ayça İrem YILMAZ

Student No.: N19131024

Department: Foreign Languages Education

Program: English Language Education Program

Status: ☒ Masters ☐ Ph.D. ☐ Integrated Ph.D.

Signature

ADVISOR APPROVAL

APPROVED
(Asst. Prof., İsmail Firat ALTAY,)

APPENDIX-G: Yayınlama ve Fikrî Mülkiyet Hakları Beyanı

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kâğıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayınlanan "**Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge**" kapsamında tezim aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H.Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- O Enstitü/Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihinden itibaren 2 yıl ertelenmiştir. ⁽¹⁾
- O Enstitü/Fakülte yönetim kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihinden itibaren ... ay ertelenmiştir. ⁽²⁾
- O Tezimle ilgili gizlilik kararı verilmiştir. ⁽³⁾

18 /07 /2022

Ayça İrem YILMAZ

"Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge"

- (1) Madde 6. 1. Lisansüstü teze ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu iki yıl süre ile tez erişime açılmasının ertelenmesine karar verebilir.
- (2) Madde 6.2. Yeni teknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internetten paylaşılması durumunda 3. şahıslara veya kurumlara haksız kazanç; imkânı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulunun gerekçeli kararı ile altı ayı aşmamak üzere tez erişime açılması engellenebilir.
- (3) Madde 7. 1. Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir*. Kurum ve kuruluşlarla yapılan işbirliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, ilgili kurum ve kuruluşun önerisi ile enstitü veya fakültenin uygun görüşü üzerine üniversite yönetim kurulu tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir.

Madde 7.2. Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir

*Tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu tarafından karar verilir.

