T.C.

TOKAT GAZİOSMANPAŞA ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

EĞİTİM BİLİMLERİ ANABİLİM DALI

İNGİLİZ DİLİ EĞİTİMİ YÜKSEK LİSANS PROGRAMI

YAPAY ZEKA SOHBET BOTLARININ EFL LİSE ÖĞRENCİLERİNİN KONUŞMA BECERİSİ, KAYGI, MOTİVASYON VE DİJİTAL OKURYAZARLIK ÜZERİNDEKİ ETKİSİNİN İNCELENMESİ

YÜKSEK LİSANS TEZİ

Mehmet TAŞDEMİR

Danışman: Dr. Öğr. Üye. Zafer SUSOY

TOKAT- 2025

ETİK SÖZLEŞME

Tokat Gaziosmanpaşa Üniversitesi Lisansüstü Eğitim Enstitüsü tez yazım kılavuzuna göre, Dr. Öğr. Üye Zafer SUSOY danışmanlığında hazırlamış olduğum "Yapay Zeka Sohbet Botlarının EFL Lise Öğrencilerinin Konuşma Becerisi, Kaygı, Motivasyon Ve Dijital Okuryazarlık Üzerindeki Etkisinin İncelenmesi" adlı Yüksek Lisans tezinin bilimsel etik değerlere ve kurallara uygun, özgün bir çalışma olduğunu, aksinin tespit edilmesi halinde her türlü yasal yaptırımı kabul edeceğimi beyan ederim.

…/…/…

Mehmet TAŞDEMİR

JÜRİ KABUL VE ONAY

Mehmet TAŞDEMİR tarafından hazırlanan "**Yapay Zeka Sohbet Botlarinin EFL Lise Öğrencilerinin Konuşma Becerisi, Kaygi, Motivasyon Ve Dijital Okuryazarlik Üzerindeki Etkisinin İncelenmesi**" adlı tez çalışmasının savunma sınavı 08/10/2025 tarihinde yapılmış olup aşağıda verilen Jüri tarafından Oy Birliği ile Tokat Gaziosmanpaşa Üniversitesi Lisansüstü Eğitim Enstitüsü İngiliz Dili Eğitimi'nde Yüksek Lisans Tezi olarak kabul edilmiştir.

Jüri Üyeleri (Unvanı, Adı Soyadı)	**İmzası**
Üye (Başkan) : Dr. Öğr. Üye Zafer SUSOY	……...
Üye : Doç. Dr. Züleyha ÜNLÜ CARLIK	…….....
Üye : Doç. Dr. Meral CEYLAN	…….....

ONAY

……/……/…….

Lisansüstü Eğitim Enstitüsü Müdürü

ABSTRACT

AN INVESTIGATION OF AL CHATBOTS ON EFL HIGH-SCHOOL LEARNERS' SPEAKING PROFICIENCY, ANXIETY, MOTIVATION AND DIGITAL LITERACY

Taşdemir, Mehmet

Master's Thesis – English Language Teaching

Advisor: Dr. Öğr. Üye Zafer SUSOY

October, 2025, x + 104 Pages

This study investigates the impact of AI-driven interaction on English as a Foreign Language (EFL) learners' speaking proficiency, speaking anxiety, motivation, and digital literacy. With the rapid advancements in artificial intelligence, AI-based conversational agents have emerged as potential tools to enhance oral communication skills in language learning. This research employs a quasi-experimental design with a mixed-methods approach, combining quantitative and qualitative data collection techniques.

The study involves 60 high school 11th-grade EFL students, divided into an experimental group (n=30) utilizing AI-powered conversational (ChatGPT and Google Gemini) for speaking practice, and a control group (n=30) engaged in traditional teacher-guided speaking activities. Participants were selected from similar classes. Over a ten-week period, the experimental group engaged in AI-driven speaking tasks, shared their recordings via Google Classroom, and received feedback from both AI and teacher. The control group continued with the regular coursebook-based speaking activities under teacher supervision.

Data collection included pre- and post-test assessments of speaking proficiency (evaluated using TOEFL iBT rubrics), speaking anxiety (measured using the FLSAS Questionnaire by Kayaoğlu & Sağlamel, 2013), and digital literacy (assessed using the scale by Hamutoğlu et al., 2017). Additionally, students' motivation levels were examined through Schmidt et al.'s (1996) language learning motivation scale. Qualitative data were obtained through semi-structured interviews with the experimental group and classroom observations.

Quantitative analysis was conducted using paired-samples and independent samples t-tests to measure within-group and between-group differences. ANOVA was employed to assess statistical significance among variables, and correlation analyses were performed to examine relationships between speaking proficiency, motivation, and anxiety. Reliability analyses (Cronbach's Alpha) ensured internal consistency of the instruments. Qualitative data were analyzed through thematic and descriptive analysis to categorize students' perceptions and classroom interactions.

Findings indicate that AI-driven speaking practice significantly improved students' speaking achievement scores compared to traditional instruction. Moreover, AI-integrated learning contributed to lower speaking anxiety and increased motivation for language learning. Qualitative data highlighted students' positive attitudes towards AI tools, particularly in receiving immediate feedback and engaging in autonomous learning. The results underscore the pedagogical value of AI-driven tools in supporting self-regulated speaking practice and enhancing digital literacy.

This study provides empirical evidence on the role of AI in language education, offering insights for integrating AI technologies into EFL curricula. Future research may examine the long-term effects of AI-based speaking interventions and their impact on learners with varying proficiency levels.

Keywords: AI-Driven Interaction, English Speaking Proficiency, EFL Learners, Speaking Anxiety, Digital Literacy, Quasi-Experimental Study.

ÖZET

YAPAY ZEKA SOHBET BOTLARININ EFL LİSE ÖĞRENCİLERİNİN KONUŞMA BECERİSİ, KAYGI, MOTİVASYON VE DİJİTAL OKURYAZARLIK ÜZERİNDEKİ ETKİSİNİN İNCELENMESİ

Taşdemir, Mehmet

Yüksek Lisans, İngiliz Dili Eğitimi Bilim Dalı

Advisor: Dr. Öğr. Üye Zafer SUSOY

Ekim, 2025, x + 104 Sayfa

Bu çalışma, yapay zeka destekli etkileşimin İngilizceyi yabancı dil olarak öğrenen (EFL) öğrencilerin konuşma becerisi, konuşma kaygısı, motivasyonu ve dijital okuryazarlıkları üzerindeki etkisini incelemektedir. Yapay zekanın hızla gelişmesiyle birlikte, yapay zeka tabanlı konuşma ajanları, dil öğreniminde sözlü iletişim becerilerini geliştirmek için potansiyel araçlar olarak ortaya çıkmıştır. Bu araştırma, nicel ve nitel veri toplama tekniklerini birleştiren karma yöntem yaklaşımıyla yarı deneysel bir desen kullanmaktadır.

Çalışma, 11. sınıfta öğrenim gören 60 lise öğrencisini kapsamaktadır. Katılımcılar, yapay zeka destekli konuşma pratiği için ChatGPT ve Google Gemini gibi araçları kullanan deney grubu (n=30) ile geleneksel öğretmen rehberliğinde konuşma etkinliklerine katılan kontrol grubu (n=30) olmak üzere ikiye ayrılmıştır. Katılımcılar aynı okuldaki sınıflardan seçilmiştir. On hafta boyunca deney grubu, yapay zeka destekli konuşma görevleri gerçekleştirmiş, bu görevleri Google Classroom üzerinden paylaşmış ve hem yapay zeka hem de öğretmenden geri bildirim almıştır. Kontrol grubu ise öğretmen rehberliğinde akran etkileşimli konuşma etkinliklerine katılmıştır.

Veri toplama sürecinde, konuşma becerisi (TOEFL iBT rubrikleri kullanılarak değerlendirilmiştir), konuşma kaygısı (Kayaoğlu ve Sağlamel, 2013 tarafından geliştirilen FLSAS Anketi ile ölçülmüştür) ve dijital okuryazarlık (Hamutoğlu vd., 2017 tarafından geliştirilen ölçekle değerlendirilmiştir) ön test ve son test yoluyla ölçülmüştür. Ayrıca, öğrencilerin motivasyon seviyeleri Schmidt vd.'nin (1996) dil öğrenme

motivasyonu ölçeği ile incelenmiştir. Nitel veriler ise deney grubundaki öğrencilerle yapılan yarı yapılandırılmış görüşmeler ve sınıf içi gözlemler yoluyla elde edilmiştir.

Nicel veriler, dağılımın normal olması nedeniyle gruplar arası ve grup içi farklılıkları belirlemek için bağımlı ve bağımsız örneklem t-testi ile analiz edilmiştir. Değişkenler arasındaki anlamlılık ANOVA testi ile değerlendirilmiş ve konuşma becerisi, motivasyon ve kaygı arasındaki ilişkileri incelemek için korelasyon analizi uygulanmıştır. Ölçeklerin iç tutarlılığını belirlemek için Cronbach's Alpha güvenirlik analizi yapılmıştır. Nitel veriler, öğrenci algıları ve sınıf içi etkileşimleri tematik ve betimsel analiz teknikleri kullanılarak incelenmiştir.

Araştırma bulguları, yapay zeka destekli konuşma pratiğinin öğrencilerin akıcılık ve sözcük çeşitliliği açısından geleneksel yönteme kıyasla önemli ölçüde gelişim sağladığını göstermektedir. Ayrıca, yapay zeka destekli öğrenmenin konuşma kaygısını azalttığı ve dil öğrenme motivasyonunu artırdığı belirlenmiştir. Nitel veriler, öğrencilerin yapay zeka araçlarına yönelik olumlu tutumlarını ortaya koymuş, özellikle anında geri bildirim alma ve bağımsız öğrenme süreçlerinde daha fazla etkileşim sağladıkları gözlemlenmiştir. Bu sonuçlar, yapay zeka destekli araçların özerk konuşma pratiğini destekleme ve dijital okuryazarlığı artırma konusundaki pedagojik değerini vurgulamaktadır.

Bu çalışma, yapay zekanın dil eğitimindeki rolüne dair ampirik kanıtlar sunarak, yapay zeka teknolojilerinin EFL müfredatına entegrasyonu için önemli çıkarımlar sunmaktadır. Gelecekteki araştırmalar, yapay zeka tabanlı konuşma uygulamalarının uzun vadeli etkilerini ve farklı yeterlilik düzeylerindeki öğrenciler üzerindeki etkilerini inceleyebilir.

Anahtar Kelimeler: Yapay Zeka Destekli Etkileşim, İngilizce Konuşma Becerisi, EFL Öğrencileri, Konuşma Kaygısı, Dijital Okuryazarlık, Yarı Deneysel Çalışma.

CONTENTS

Page

ETIK SÖZLEŞME i

JURI KABUL VE ONAY ii

ABSTRACT iii

ÖZET v

LIST OF ABBREVIATIONS AND SYMBOLS x

INTRODUCTION 1

- Problem of the Study 1
- Research Aim 2
- Significance of the Study 3
- Assumptions of the Study 5
- Limitations 5
- Definitions 6

CONCEPTUAL FRAMEWORK 9

METHOD 13

- Research Design 13
- Sample of the Study 13
- Data Collection Tools 14
- Reliability of the Speaking Performance Criteria (Pre-Test) 14
- Reliability of the Speaking Performance Criteria (Post-Test) 15
- Reliability of the Speaking Anxiety Scale (Pre-Test) 17
- Reliability of the Speaking Anxiety Scale (Post-Test) 18
- Reliability of the Motivation Scale (Pre-Test) 19
- Reliability of the Motivation Scale (Post-Test) 20
- Reliability of the Digital Literacy Scale (Pre-Test) 22
- Reliability of the Digital Literacy Scale (Post-Test) 23
- Data Collection Produce 24
- Data Analysis 26

RESULTS 28
Quantitative Results 28
Digital Literacy 31
The paired samples t-test of Digital Literacy 33
Independent Sample T-Tests of Digital Literacy Skills Between Groups 35
Pre-test Results 35
Post-test Results 36
Item-Level Analysis of the Digital Literacy Scale 37
FLSAS (Foreign Language Speaking Anxiety Scale) scores. 39
Paired Samples T-tests 41
Independent Sample t-tests 43
Item-Level Analysis of (FLSAS) 45
Foreign Language Learning Motivation 47
Paired Sample T-Test 49
Independent Samples T-test 51
Item-Level Analysis of the Foreign Language Motivation Scale 53
Speaking Performance 55
Paired Samples T-Test 56
Independent Samples T-test 58
Qualitative Results 60
DISCUSSION 65
CONCLUSION & RECOMMENDATIONS 68
REFERENCES 70
APPENDIX 74
Appendix 1. Digital Literacy Scale 74
Appendix 2 Foreign Language Motivation Scale 76
Appendix 3 Foreign Language Speaking Anxiety Scale 82
Appendix 4 Semi Structured Interview Questions 85
Appendix 5 TOEFL IBT Independent Speaking Rubric 86
Appendix 6 Item Analysis of Motivation Scale Pre-Test Results 87
Appendix 7 Item Analysis of Motivation Scale Post-Test Results 91

Appendix 8 Item Analysis of Digital Literacy Scale Pre-Test Results 95
Appendix 9 Item Analysis of Digital Literacy Scale Post-Test Results 96
Appendix 10 Item Analysis of Findings of FLSAS Scale Pre-Test Results 97
Appendix 11 Item Analysis of Findings of FLSAS Scale Post-Test Results 99
Appendix 12 Approval of Ministry of Education 101
Appendix 13 Institute Board of Directors' Decision 102
Appendix 14 Ethics Committee Decision 103
ÖZGEÇMİŞ 104

LIST OF ABBREVIATIONS AND SYMBOLS

AI : Artificial Intelligence (Yapay Zekâ)

ANOVA : Analysis of Variance (Varyans Analizi)

EFL : English as a Foreign Language (Yabancı Dil Olarak İngilizce)

FLSAS : Foreign Language Speaking Anxiety Scale (Yabancı Dil Konuşma Kaygısı Ölçeği)

FL : Foreign Language (Yabancı Dil)

M : Mean (Ortalama)

SD : Standard Deviation (Standart Sapma)

n : Number of Participants (Katılımcı Sayısı)

TOEFL iBT : Test of English as a Foreign Language – Internet Based Test

T-test : Student's t-test (İstatistiksel Test)

df : Degrees of Freedom (Serbestlik Derecesi)

p : Probability Value (Anlamlılık Düzeyi)

r : Correlation Coefficient (Korelasyon Katsayısı)

α (Cronbach's Alpha) : Güvenirlik Katsayısı

INTRODUCTION

Problem of the Study

Speaking, as one of the core components of language proficiency, poses a unique challenge for EFL learners, particularly in secondary education contexts such as Turkish high schools (Celik et al., 2025). Unlike receptive skills like reading and listening, speaking demands real-time processing, active language production, and interaction, often under the psychological pressure of being evaluated by others. In many traditional language classrooms, speaking activities are limited both in time and depth, often reduced to short, scripted dialogues or rehearsed presentations that fail to foster genuine communication or fluency (Shahriar, 2025). As a result, students are left without adequate opportunities to develop the spontaneous and interactive language use that is central to communicative competence.

Adding to the situation is that speaking a foreign language is quite stressing among students. Many learners feel an extraordinary amount of fear and discomfort when asked to speak in a foreign language and are particularly sensitive to doing so in front of peers or a teacher. Speaking anxiety can lead students to try to withdraw completely, avoid speaking, and participate very little. These emotional barriers create a vicious cycle in which students speak less, receive less feedback, and progress more slowly (Baskara, 2025). Furthermore, teachers in a traditional classroom setting often lack time and resources needed to provide anything close to individualized instruction and assistance to help their students overcome these affective barriers.

Moreover, the issue of motivation in language learning further complicates the development of speaking skills. When learners do not see immediate progress or fail to engage with meaningful and relevant content, their intrinsic motivation can diminish (Han, 2020). Traditional classroom settings may not always provide the personalized, engaging, and interactive environments necessary to sustain learners' interest and commitment—particularly when language learning is perceived as abstract or disconnected from real-life communication.

Additionally, as the world becomes increasingly digital, students are expected to integrate technological tools into their learning routines. Yet, there exists a gap between

the potential of technology in language education and its actual implementation in classrooms (Al Mahrooqi, 2024). Many students, despite being digital natives, lack the specific digital literacy required to effectively use educational technologies for language learning purposes. Meanwhile, many schools remain slow to adopt innovative digital tools, often due to limited resources, lack of teacher training, or curricular constraints.

The emergence of new artificial intelligence (AI) applications, particularly AI-enabled conversational bots or chatbots, represents another promising avenue for language learning. These tools are intended to provide students with an engaging, interactive, and low-stakes speaking environment (Al Mahrooqi, 2024). Nevertheless, while it is possible to argue theoretically that AI applications can address some of the challenges in EFL such as lack of practice time, anxiety around speaking, motivation, and digital inequalities, empirical studies have not kept pace. In general, research studies conducted in university or adult education contexts have primarily studied writing and reading (not speaking) and have focused on the impact of AI applications on writing or reading only. Research that considers the emotional, cognitive, and behavioral impacts of AI speaking practice for students in secondary education contexts (in mainstreaming learning time or outside of school) is still very new (Yang et al., 2022).

The absence of a comprehensive, school-level, data-driven empirical investigation of AI-supported speaking tools has created a knowledge void within language education. When we lack a comprehensive understanding of employing such tools to influence speaking performance, learner anxiety, motivation, and digital competence, educators and policy makers are unable to draw on the knowledge to make informed decisions about the use of AI in the language classroom (Zhou et al., 2025). Consequently, examining the challenges related to speaking instruction in traditional contexts and new opportunities with AI is necessary to further inform future pedagogies and identify deficiencies within foreign language instruction.

Research Aim

The primary aim of this study is to investigate the effects of artificial intelligence (AI)-supported chatbot interaction on the speaking performance, speaking anxiety, motivation and digital literacy of Turkish high school students learning English as a

foreign language (EFL). As speaking remains a particularly underdeveloped and anxiety-inducing skill in many traditional language learning environments, this research seeks to determine whether AI tools such as ChatGPT and Google Gemini can offer a meaningful and pedagogically sound alternative to conventional classroom practices. The study is designed to explore not only the linguistic outcomes of AI-supported speaking activities but also their impact on learners' emotional responses and levels of engagement.

Through a mixed-methods research design that applies a quasi-experimental design, this research seeks to gather quantitative and qualitative data to provide a holistic understanding of how future AI-based speaking tasks impact learners' speaking achievement outcomes, speaking anxiety, motivation, and digital literacy. While quantitative tools will measure levels of change of speaking performance, anxiety, motivation and digital literacy, qualitative tools such as interviews and observations will give a deeper insight of students' personal experiences of AI-based speaking practice.

Another key aim of this research is to analyze the relationships between speaking skills, anxiety, motivation, and digital literacy within the context of AI-supported instruction. The study intends to identify patterns and correlations that can inform future teaching practices and the integration of technology in EFL classrooms. Furthermore, by comparing the outcomes of AI-supported and traditional speaking activities, the research aims to highlight the strengths and limitations of both approaches, offering a comparative perspective that can guide educators in selecting appropriate methods for different learner profiles.

In conclusion, this research aims to add to the scholarship surrounding the use of AI in language education by providing empirical data regarding high school EFL learners, one of the under-researched areas in this field. In doing so, this work hopes to contribute to more personalized, engaging, and effective speaking instruction by providing frameworks, guidelines, and recommendations related to the use of new technologies and the modes and methods of assessment that underpin research in technology studies.

Significance of the Study

This research is relevant and significant to several communities in the field of English language education, including researchers, teachers, curriculum designers, and

education policymakers. As we become more and more integrated with artificial intelligence in educational contexts, it is essential to understand how these tools can be used purposively and effectively for language learning (Hapsari & Wu, 2022). While this study is specifically focused on the effects of AI-mediated chat tool interaction on aspects of speaking abilities, anxiety and motivation, it provides relevant research to inform pedagogical practices as classrooms continue to evolve technologically.

For applied linguistics and language education, the investigation presents quantitative and qualitative evidence on the pragmatic use of AI tools to assist one of the most demanding areas of language learning, namely speaking. It contributes to a scientific understanding of the extent to which digital tools may enhance learning environments that are supportive and engaging, especially for students who may be reluctant or anxious to participate in traditional classroom exchanges. In this sense, the study fills a key gap in the literature by focusing on high school learners, which is a group that has been under-explored in experimental AI-related language research, which biases towards working with adults or university students.

From an instructional perspective, the findings are expected to offer valuable guidance to English teachers seeking alternative methods for enhancing oral communication skills among learners (Ma et al., 2024). The integration of AI tools into classroom practice has the potential to personalize instruction, foster learner autonomy, and reduce classroom anxiety, thereby supporting more inclusive and differentiated learning experiences. Teachers can benefit from evidence-based insights on how to adapt or complement traditional speaking activities with AI-mediated tasks to improve student outcomes.

For curriculum developers and educational leaders, the study's results may inform future policies regarding the integration of educational technology into foreign language teaching programs (Ballıdağ & Aydın, 2025). As national and institutional education strategies increasingly prioritize digital transformation, research-driven recommendations on how AI tools can be effectively and ethically used in secondary education will be crucial in guiding decisions related to resource allocation, teacher training, and curriculum planning.

Overall, this study aspires to contribute to the ongoing discourse on educational innovation by providing a realistic and evidence-based perspective on how AI-supported language learning can be used not to replace, but to enhance, existing pedagogical practices. It emphasizes the role of technology not as a standalone solution, but as a meaningful supplement that can address some of the persistent challenges in foreign language instruction—particularly in improving speaking competence, reducing affective barriers, and increasing student motivation in the modern classroom.

Assumptions of the Study

In the design and implementation of this research, several key assumptions were established to ensure the integrity and relevance of the study. First, it was presumed that the measurement tools employed specifically the speaking performance rubric, speaking anxiety questionnaire, motivation scale, and digital literacy scale would effectively gather pertinent data that aligns with the variables and objectives of the research.

Additionally, it was assumed that the expert feedback received during the development and adaptation of the interview questions and observation protocols was very thorough to support content relevance and clarity for productive analysis. The level of participation was also assumed to be similar because all participants in both groups (experimental and control) were anticipated to regularly attend the speaking practice sessions and complete the planned activities during the research period.

Moreover, it was assumed that the classroom context, lesson time, and general curriculum for both groups would be consistent throughout the 10-week implementation, in all respects apart from the experimental treatment. Finally, it was assumed that technology (both digital tools and AI platforms) used by the study - including ChatGPT, Google Gemini, and Google Classroom - would work as designed and as intended, allowing students in the experimental group to participate solve fully engage in the AI supported speaking tasks aimed to enhance their learning experience. The study itself rests on these assumptions and is predicated on a set of useful assumptions that are vitally important to some findings of the research.

Limitations

The current research has certain limitations that may influence the generalizability and applicability of the findings. First, it focused on just 11th-grade students in one public high school in Turkey in the 2024–2025 academic year, which limited the sample to one age group and one context of education. Second, a relatively small sample of 60 students, with equal numbers of students in each experimental and control group, curtails the ability to generalize findings to other grade levels, or other regional contexts in Turkey that are diverse. Third, the duration of the study (10 weeks) likely was not adequate to capture an adequate understanding of long-term effects on learning outcomes in which students have been using AI-supported speaking practice. Fourth, this research only promoted a limited number of AI tools (ChatGPT, Gemini, and Google Classroom) that did not include other cutting-edge AI tools that would likely add value to the study of student engagement and student performance. Finally, since the findings were related to the participants in the context of their perceptions and experiences in this school, they are not generalizable to all Turkish high school English as a Foreign Language (EFL) learners, because different contextual factors (e.g. previous experience, student characteristics) will contribute to the same study in a different context.

Definitions

Artificial Intelligence (AI):

In this research, artificial intelligence is the use of computer-based technologies that can usually perform tasks that require human intelligence, such as understanding and generating language. More specifically, AI is represented in this research by chatbot tools--ChatGPT and Google Gemini--that EFL learners can use to engage in interactive speaking practice, such as collaborating on a text conversation.

AI-Supported Chatbot Interaction:

This pertains to the participation of learners in conversations with AI chatbots designed to mimic human conversations and could be text- or voice-based. For this study, "conversing" will refer to weekly speaking activities conducted via the chatbots ChatGPT

and Google Gemini, along with an accompanying feedback process using Google Classroom.

Speaking Skills:

Speaking skills denote the learners' abilities to communicate orally in English, which encompasses fluency, pronunciation, grammatical accuracy, vocabulary, and coherent. In the present study speaking skills are assessed by utilizing the assessment criteria based on TOEFL iBT speaking rubrics.

Speaking Anxiety:

Speaking anxiety refers to worry or fear felt by learners when communicating verbally in the foreign language. For the present study, it will be measured with the Foreign Language Speaking Anxiety Scale (FLSAS), and is defined as an affective factor that can influence learners' willingness to engage in speaking activities (Kayaoğlu & Sağlamel, 2013).

Motivation:

In the study, motivation is defined as both the external and internal factors that influence learners' willingness and effort to improve their English-speaking skills. Motivation was measured with a language learning motivation scale adapted from Schmidt (1996), which examined both intrinsic and extrinsic motivation component.

Digital Literacy:

Digital literacy encompasses the capacity to proficiently utilize digital technologies to locate, evaluate, generate, and communicate digital information. Within the context of this study, digital literacy refers specifically to the students' ability to use AI-enabled educational tools and platforms as part of their learning process (for example, chatbots and Google Classroom) and is measured via a digital literacy scale developed by Hamutoğlu et al. (2017).

Traditional Speaking Practice:

This designation is given to speaking activities taking place in the classroom without AI. In these activities, the teacher is generally in front of the class, leading instruction or interacting with peers in activities such as pair work, group discussions, and

group role-plays. These interactions take place, in some form, in a typical classroom context by a trained teacher.

CONCEPTUAL FRAMEWORK

The conceptual framework of this study is grounded in the intersection of language learning theory, affective factors in second language acquisition, motivation theory, digital literacy, and the pedagogical integration of artificial intelligence (AI) in education. This framework establishes the foundation upon which the study is designed and implemented, offering a structured lens through which the relationships among the variables—speaking skills, speaking anxiety, motivation, and digital literacy—can be explored in the context of AI-supported chatbot interaction.

Speaking Skills in Foreign Language Learning

Although speaking is generally considered the most complex and demanding of the four language skills, it requires not only linguistic competence, but also cognitive processing, fluency, sociolinguistic knowledge, and confidence. In EFL situations, particularly in typical high school contexts, learners generally lack authentic communication and opportunities to practice speaking spontaneously. Instruction is typically teacher-centered, and most speaking tasks in classrooms are reduced to controlled, low-variability activities that do not provide any reflection of real-life communication that an individual might engage in. There is a variety of theory related to language learning; for example, Krashen's Input Hypothesis (1985) argues that language acquisition happens when learners are provided with comprehensible input slightly above their currently level (i+1). Input is necessary, but it is not a sufficient condition for developing the productive skills, such as speaking. Long's Interaction Hypothesis (1996) builds upon this intuitive notion by stating that interaction and negotiation of meaning are critical to language learning. The Interaction Hypothesis asserts that output, along with feedback provided during communicative interaction, allows for interlanguage development. In this way, AI chatbots provide an attractive option for creating meaningful interaction in a safe and low-threat digital space. Continuous and contextualized practice may improve aspects of speaking performance, for example fluency, grammatical accuracy, pronunciation and coherence.

Speaking Anxiety and Affective Factors

Affective factors, particularly foreign language anxiety (FLA), can play a significant role in understanding an individual's ability to acquire a new language. Horwitz et al 1986) define language anxiety as a context-specific type of anxiety, distinct from anxiety in a general way. Language anxiety involves nervousness and apprehension related to the individual using a foreign language. Where speaking is concerned, FLA may often arise from fearing negative evaluation, communication apprehension, and/or low self-efficacy; and particularly in classroom contexts where learners are observed by their peers and teachers.

This question can be examined through Vygotsky's Sociocultural Theory (1978) when he emphasized a belief that learning cannot be divorced from its social mediation. Learning, according to Vygotsky, is only possible in the presence of a more capable other, within an individual's Zone of Proximal Development (ZPD). The 'more capable other' in traditional contexts may be a teacher or peer; now AI tools may serve this role in some contexts, as well. Within certain contexts, AI tools may provide the same scaffolded interaction, support, and feedback as peer interactions, without the weight of social accountability. Operating from this theoretical perspective, AI-supported speaking engagement spaces can offer learners a stress-free space to practice speaking, make mistakes, and receive feedback because there is a decreased affective filter without the burden of social evaluation, resulting in increased willingness to participate.

Motivation in Language Learning

Motivation is a key determinant of success in foreign language learning. Gardner's (1972) socio-educational model identifies two primary forms of motivation: integrative motivation, which involves a desire to integrate into the target language culture, and instrumental motivation, which is oriented toward practical goals such as passing exams or securing employment. Meanwhile, Deci and Ryan's (1985) Self-Determination Theory distinguish between intrinsic motivation (engaging in an activity for its inherent satisfaction) and extrinsic motivation (performing a task for an external reward or outcome).

AI tools have the potential to foster both forms of motivation. Intrinsically, learners may enjoy the novelty, interactivity, and immediate feedback offered by chatbot tools, which can make the learning experience more engaging and autonomous. Extrinsically, AI-supported platforms may help students see clear progress and feel better prepared for exams, thereby reinforcing task value and effort. Furthermore, features such as personalized prompts, adaptive difficulty levels, and gamified elements can sustain learner interest and promote persistence in oral practice.

Digital Literacy in Modern Education

Digital literacy is defined as an ability to access, evaluate, create, and communicate information using digital technologies. Digital literacy encompasses more than basic computer skills; instead, it involves critical thinking and digital communication as well as the responsible use of technology. In the context of language education, digital literacy can have an important effect on the way learners interact with online tools and platforms (Zhang et al., 2024). Specifically, those individuals who are more digitally literate tend to be more comfortable and confident in using educational technology, which may promote positive learning outcomes.

In this study, digital literacy is both a variable of interest as well as a contextual factor which can influence AI-supported speaking practice. For example, learners' comfort and familiarity with technology can shape how quickly they adapt to a chatbot, how effectively they engage with a digital platform, and how they pursue their own learning beyond the classroom (Apriani et al., 2024). Furthermore, as AI is incorporated into educational policy and curricular design, promoting digital literacy is essential for equitable access to and implementation of new tools.

AI-Supported Chatbot Interaction

There has been a rapid proliferation of AI chatbots such as ChatGPT and Google Gemini that replicate human-like conversation and are finding applications in the field of education for personalizing and humanizing learning experiences. For language learning, this technology can function as a conversation partner that is easily accessible, unconditionally patient, and notably utilitarian (Wei & Zhao, 2025). Language learners

are able to practice self-paced low-pressure conversations, while obtaining immediate feedback on their language purportedly in context. A language learner can retrieve contextualized practice, lexically advance, and receive corrective feedback free from the worry of being judged.

AI tools are also generally aligned with Task-Based Language Teaching (TBLT), with TBLT emphasizing communicating meaning through tasks that reflect real-world language use. AI chatbots can be designed to use chatbots against the backdrop of authentic communication tasks such as ordering food, interviewing for a job, and voicing opinions related to modern happenings (Iswari et al., 2024). Ultimately, this can lead to communicative competence while utilizing request-based principles that can relate to real-life language functions.

Integration of Theory and Conceptual Model

The collected data are subjected to a rigorous analysis employing both statistical and qualitative techniques to ensure a comprehensive understanding of the findings. For quantitative data analysis, several methods are utilized, beginning with the Paired Samples t-Test to evaluate within-group changes between pre-test and post-test scores for both the experimental and control groups. Additionally, an Independent Samples t-Test is conducted to discern performance differences between these two groups, while Correlation Analysis is employed to investigate the relationships among key variables such as speaking skills, speaking anxiety, motivation, and digital literacy. To ensure the reliability of the scales used in the study, a Reliability Analysis using Cronbach's Alpha is performed, verifying the internal consistency of the measures. On the qualitative side, Thematic Analysis is applied to interview transcripts to uncover recurring themes and patterns that shed light on students' experiences with AI chatbots. Complementing this, Descriptive Analysis of observation data provides a nuanced interpretation of student behaviors, engagement levels, and classroom dynamics. Through this multifaceted methodology, the study aims to deliver evidence-based insights into the pedagogical benefits of utilizing AI tools in enhancing speaking development within English as a Foreign Language (EFL) classroom.

METHOD

This section outlines the research design, participants, data collection tools, procedures, and data analysis techniques used in the study. The methodology has been carefully structured to explore the effects of AI-supported chatbot interaction on the FL speaking achievement scores, FL speaking anxiety, FL speaking motivation and digital literacy of Turkish high school students learning English as a foreign language.

Research Design

This study adopts a quasi-experimental mixed-methods design (Campbell 1963), which combines both quantitative (Fisher et al., 2014) and qualitative Erickson (2023) research methods. The quasi-experimental structure includes pre-tests and post-tests administered to both an experimental group and a control group. The integration of qualitative data through interviews and observations supports the quantitative findings and provides deeper insight into students' experiences with AI-supported speaking practice. This design allows for a holistic evaluation of the impact of AI tools on language learning outcomes and affective factors.

Sample of the Study

This study involves 60 Turkish high school students from the 11th grade participating during the 2024–2025 academic year. Participants were selected through convenient sampling, as the researcher had direct access to the school and classes where the implementation was carried out. The students are divided into two distinct groups to evaluate the effectiveness of AI-supported learning methodologies. The Experimental Group, consisting of 30 students, engages in innovative learning through weekly AI-supported speaking activities that utilize advanced tools like ChatGPT and Google Gemini. These students collaborate on sharing their interactions with the chatbots via Google Classroom, fostering an interactive learning environment where they receive valuable feedback from both their peers and teachers. In contrast, the Control Group, also made up of 30 students, partakes in traditional teacher-led activities that focus on developing speaking skills through methods such as pair and group work, role plays, and engaging in classroom discussions, but without the integration of any AI tools. To

maintain consistency in the study, all participants are drawn from classrooms with similar English proficiency levels, ensuring a homogeneous group for accurate results.

Data Collection Tools

In order to thoroughly investigate the research questions, a combination of both quantitative and qualitative data collection instruments was utilized. The Speaking Skills Assessment involves pre-test and post-test speaking tasks, which are evaluated using a rubric adapted from the TOEFL iBT Speaking Scoring Guide which can be seen in Appendix 5, focusing on key dimensions such as fluency, pronunciation, grammar, vocabulary, and coherence.

Reliability of the Speaking Performance Criteria (Pre-Test)

To evaluate the internal consistency of the speaking performance criteria prior to the intervention, a reliability analysis was conducted using Cronbach's Alpha. The analysis included all 60 valid cases, and assessed the four components of speaking performance: General Description, Delivery, Language Use, and Topic Development, based on a rubric adapted from the TOEFL iBT Speaking Scoring Guides seen in Table 1 the analysis yielded a Cronbach's Alpha coefficient of .851, with a standardized alpha of .852, indicating a high level of internal consistency across the four rubric dimensions (Steiner, 2003).

Table 1. Reliability Statistics of Pre-Test Performance Criteria

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N
.851	.852	4

The mean inter-item correlation in Table 2 was calculated as .590, with individual correlations ranging from .491 to .656, reflecting strong internal alignment among the components of the rubric.

Table 2. Summary Item Statistics of Pre-Test Speaking Performance Criteria

	M	Min	Max	Range	Max / Min	Variance	N
Inter-Item Correlations	.590	.491	.656	.165	1.335	.003	4

Further examination of the item-total statistics in Table 3 showed that all criteria contributed meaningfully to the overall reliability. Notably, the "Topic Development" item demonstrated the strongest corrected item-total correlation (.753), followed by "Language Use" (.715), both of which suggest a particularly strong alignment with the overall scale (Steiner, 2003). The removal of any individual item would reduce the overall alpha, further supporting the cohesiveness of the selected rubric dimensions.

Table 3. Item-Total Statistics of Pre-Test Speaking Performance Criteria

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
General Description	33.3500	127.858	.642	.428	.832
Delivery	33.9000	124.769	.660	.454	.824
Language Use	33.0667	127.012	.715	.515	.801
Topic Development	33.8833	123.461	.753	.567	.785

These results confirm the internal reliability of the speaking performance assessment tool for pre-test evaluations and justify its application in measuring learners' oral proficiency in the context of the study.

Reliability of the Speaking Performance Criteria (Post-Test)

Following the intervention phase, the reliability of the speaking performance assessment criteria was reanalyzed using Cronbach's Alpha. The post-test analysis, conducted on 60 valid cases, included the same four evaluative dimensions utilized in the pre-test: General Description, Delivery, Language Use, and Topic Development, based on the adapted TOEFL iBT Speaking Rubric.

The reliability analysis in Table 4 produced a Cronbach's Alpha coefficient of .932, with a standardized alpha of .933, reflecting excellent internal consistency among the assessment criteria (Steiner, 2003). This substantial increase from the pre-test alpha (.851) suggests that the rubric items functioned with even greater internal coherence after the instructional intervention, possibly due to more consistent performance patterns among students.

Table 4. Reliability Statistics of Post-test Speaking Performance Criteria

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.932	.933	4

Inter-item correlations in Table 5 were notably high, with a mean of .776 and a range from .725 to .797, indicating strong alignment across all scoring dimensions. Item-total statistics further confirmed that each component contributed significantly to the overall reliability.

Table 5. Summary Item Statistics of Post-test Speaking Performance Criteria

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Inter-Item Correlations	.776	.725	.797	.072	1.100	.001	4

The corrected item-total correlations in Table 6 ranged from .812 (General Description) to .866 (Topic Development), while the alpha values if each item were deleted remained consistently below the total alpha, confirming that all items strengthened the scale's internal consistency.

Table 6. Item-Total Statistics of Post-test Speaking Performance Criteria

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
General Description	41.0333	133.592	.812	.668	.921
Delivery	40.8167	132.729	.849	.725	.909
Language Use	40.7000	134.315	.838	.714	.913
Topics Development	41.3000	126.654	.866	.750	.904

These findings demonstrate that the speaking rubric was not only reliable at the outset of the study but also increasingly consistent in assessing post-intervention speaking performances. This further supports its validity as a robust instrument for evaluating oral proficiency gains in AI-supported language learning contexts.

To gauge students' levels of anxiety related to English speaking, the Foreign Language Speaking Anxiety Scale (FLSAS), which was translated into Turkish by Sağlamel and Kayaoğlu (2013), is employed.

Reliability of the Speaking Anxiety Scale (Pre-Test)

To evaluate the internal consistency of the Foreign Language Speaking Anxiety Scale (FLSAS) prior to the intervention, a reliability analysis was conducted using Cronbach's Alpha. The analysis included 60 valid cases, with no missing data, and encompassed all 31 items of the scale. In Table 7 the Cronbach's Alpha coefficient was calculated as .814, with a standardized alpha of .810, indicating a high level of internal consistency (Steiner, 2003).

Table 7. Reliability Statistics of Pre-Test Speaking Anxiety Scale

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.814	.810	31

In Table 8 the mean inter-item correlation was .121, with individual correlations ranging from –.287 to .503, suggesting an overall positive relationship among the scale items, despite a few weak or negatively correlated pairs. These results demonstrate that the FLSAS is a reliable instrument for measuring students' speaking anxiety in the pre-intervention phase of this study.

Table 8. Summary Item Statistics of Pre-Test Speaking Anxiety Scale

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Inter-Item Correlations	.121	-.287	.503	.790	-1.749	.017	31

Given the scale's strong internal reliability, it is considered appropriate for assessing anxiety-related affective factors in the context of AI-supported speaking practice in EFL high school classrooms.

Reliability of the Speaking Anxiety Scale (Post-Test)

The internal consistency of the Foreign Language Speaking Anxiety Scale (FLSAS) was reevaluated after the intervention using Cronbach's Alpha. The analysis was conducted on responses from 60 valid cases, encompassing all 31 items of the scale. In Table 9 the resulting Cronbach's Alpha coefficient was .886, with a standardized alpha of .885, indicating a very high level of internal reliability (Steiner, 2003).

Table 9. Reliability Statistics of Post-test Speaking Anxiety Scale

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.886	.885	31

In Table 10 the mean inter-item correlation was .200, with values ranging from –.144 to .570, reflecting generally strong and coherent relationships among the scale items.

This marked improvement from the pre-test reliability (.814) suggests increased consistency in students' responses following the AI-supported speaking intervention.

Table 10. Summary Item Statistics of Post-test Speaking Anxiety Scale

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Inter-Item Correlations	.200	-.144	.570	.713	-3.969	.014	31

The results confirm that the FLSAS remained a psychometrically sound instrument for assessing students' speaking anxiety levels in the post-intervention phase and support its continued use in evaluating affective outcomes in technology-enhanced language learning environments. Additionally, to measure students' intrinsic and extrinsic motivation for learning English, a motivation scale adapted from Schmidt (1996) is administered.

Reliability of the Motivation Scale (Pre-Test)

To ensure the internal consistency of the Motivation Scale used in the pre-test phase, a Cronbach's Alpha reliability analysis was conducted. The scale comprised a total of 72 items, administered to 60 participants. According to Table 11, the Cronbach's Alpha coefficient was calculated as .888, indicating a high level of internal consistency among the items. This suggests that the scale is a reliable instrument for measuring students' language learning motivation within the scope of the study

Table 11. Reliability Statistics of Pre-test Motivation Scale

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.888	.885	72

The analysis confirmed that none of the items needed to be removed, as the reliability remained robust across all items. Therefore, the scale was deemed suitable for

further statistical analyses related to the motivation construct in the context of AI-supported language learning interventions.

In addition to the overall reliability coefficient, Table 12 presents the summary item statistics, offering deeper insight into the inter-item correlations among the 72 items in the scale. The mean inter-item correlation was found to be .096, which, although relatively low, is acceptable for broad constructs like motivation that include diverse components. The minimum correlation observed between item pairs was –.356, while the maximum correlation reached .554, indicating a wide range of relationships between individual items.

Table 12. Summary Item Statistics of Pre-test Motivation Scale

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Inter-Item Correlations	.096	-.356	.554	.909	-1.557	.019	72

The range of inter-item correlations (.909) further reflects this variability, while the variance (.019) indicates moderate dispersion around the mean. A negative minimum value suggests that a few items may tap into different subdimensions of motivation or exhibit low alignment with the overall construct. However, since the Cronbach's Alpha remained well within acceptable limits, these variations do not compromise the reliability of the scale.

These findings underscore the multidimensional nature of the motivation construct but affirm that the scale items work coherently as a whole. The breadth of correlations is consistent with instruments designed to capture complex psychological traits such as language learning motivation.

Reliability of the Motivation Scale (Post-Test)

Following the intervention, the internal consistency of the Motivation Scale was re-evaluated using Cronbach's Alpha. The post-test data, collected from 60 valid cases with no exclusions, were subjected to reliability analysis. In Table 13 the Motivation

Scale, consisting of 72 items, yielded a Cronbach's Alpha coefficient of .904, indicating excellent internal reliability. The value based on standardized items was similarly high at .900, further supporting the internal coherence of the scale items.

Table 13. Reliability Statistics of Post-test Motivation Scale

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.904	.900	72

The inter-item correlation statistics in Table 14 revealed a mean correlation of .111, with values ranging from –.430 to .562. While a few items showed low or even negative correlations with others, the overall reliability remained high, suggesting that the scale as a whole functioned effectively in capturing the construct of motivation.

These results confirm the suitability of the Motivation Scale for post-intervention analysis and support its continued use in examining the motivational effects of AI-supported speaking practice among EFL high school students.

Table 14. Summary Item Statistics of Post-test Motivation Scale

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Inter-Item Correlations	.111	-.430	.562	.993	-1.307	.021	72

Furthermore, the digital literacy levels of the students are assessed using a scale developed by Hamutoğlu et al. (2017), which specifically examines their proficiency in using digital tools for educational purposes.

Reliability of the Digital Literacy Scale (Pre-Test)

To assess the internal consistency of the Digital Literacy Scale prior to the intervention, a reliability analysis was conducted using Cronbach's Alpha. The analysis included responses from 60 valid cases, with all 17 items of the scale considered. In Table 15 the Cronbach's Alpha coefficient was calculated as .550, with a standardized item alpha of .546, indicating a moderate level of internal consistency.

Table 15. Reliability Statistics of Pre-test Digital Literacy Scale

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.550	.546	17

In Table 16 the inter-item correlation analysis revealed a mean of .066, with correlations ranging from –.309 to .434, suggesting that some items were only weakly correlated and a few negatively correlated with others. Although the scale demonstrates modest internal reliability, it remains within an acceptable threshold for exploratory educational research, particularly in studies involving complex, multi-dimensional constructs such as digital literacy.

Table 16. Summary Item Statistics of Pre-test Digital Literacy Scale

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Inter-Item Correlations	.066	-.309	.434	.743	-1.405	.020	17

Given the moderate reliability score, the scale was retained for subsequent analysis; however, the findings derived from this instrument should be interpreted with caution. Further refinement or item-level review may enhance the scale's internal coherence in future applications.

Reliability of the Digital Literacy Scale (Post-Test)

The internal consistency of the Digital Literacy Scale was reassessed following the intervention using Cronbach's Alpha. The analysis included all 60 participants, with no excluded cases, and incorporated all 17 items on the scale. The resulting Cronbach's Alpha coefficient was .831, indicating a high level of internal consistency and a marked improvement compared to the pre-test reliability score. In Table 17 the coefficient based on standardized items was identical at .831, confirming the overall reliability of the scale after the intervention.

Table 17. Reliability Statistics of Post-test Digital Literacy Scale

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.831	.831	17

In Table 18 Inter-item correlations ranged from –.021 to .491, with a mean value of .224, suggesting a generally positive and acceptable level of item-to-item association. This improvement in internal reliability following the 10-week AI-supported instructional period may indicate enhanced coherence in students' self-reported digital literacy perceptions, likely influenced by their consistent engagement with digital tools such as AI chatbots and online platforms.

Table 18. Summary Item Statistics of Post-test Digital Literacy Scale

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Inter-Item Correlations	.224	-.021	.491	.512	-23.665	.012	17

The findings support the use of the Digital Literacy Scale in the post-test phase and strengthen its role in evaluating the impact of AI-mediated language learning environments on learners' technological competencies.

Qualitative insights are gathered through semi-structured interviews with selected students from the experimental group, allowing for an in-depth exploration of their experiences, perceptions, and attitudes toward AI-supported speaking practice. The interview protocol was developed based on a review of relevant literature on language learning anxiety, AI integration in education, and learner autonomy. To ensure the relevance and clarity of the interview questions, the researcher consulted three experts in the fields of English Language Teaching and Educational Technologies. These experts reviewed the draft questions and provided feedback on content coverage, linguistic appropriateness, and the potential of each question to elicit rich and reflective responses. After revisions based on this expert input, a pilot interview was conducted with two students outside the research sample to further validate the structure and flow of the questions. The finalized interviews were then conducted in Turkish to ensure clarity and comfort for participants and later translated into English for thematic analysis. Each session was audio-recorded with participant consent and lasted approximately 15–20 minutes, ensuring a comfortable and open atmosphere to encourage authentic responses.

Lastly, observation reports were used to document classroom dynamics, capturing student engagement, participation, and any visible signs of speaking anxiety during speaking activities. These reports offered supplementary data to triangulate findings from the other instruments and provided a comprehensive overview of the learning environment throughout the study.

Data Collection Produce

The research spans a duration of 10 weeks, meticulously structured into distinct phases to ensure comprehensive data collection and analysis. In the first week, participants undergo pre-test administration, during which they complete a series of assessments, including a speaking test and scales measuring anxiety, motivation, and digital literacy. This baseline data establishes the starting point for evaluation. During weeks two through nine, the intervention phase occurs, where the experimental group participates weekly in AI-supported speaking activities using platforms such as ChatGPT and Google Gemini. The experiment group followed out the instructions for every week

through the collaborative panel we have created in Google Classroom WEB tool as seen in Figure 1.

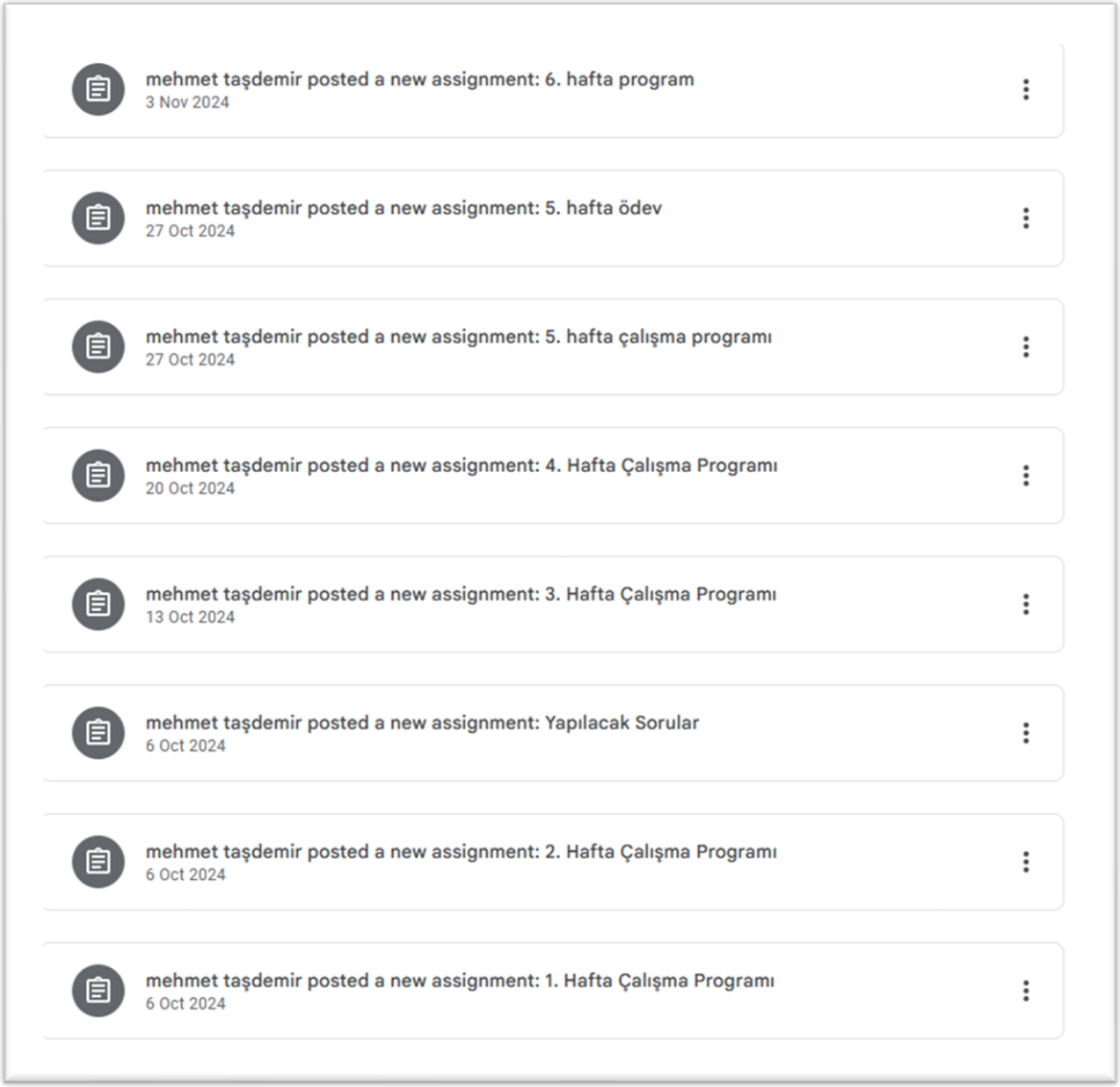


Figure 1 Showing How The Weekly Study Program Was Distributed In Google Classroom

Their interactions and feedback were facilitated through Google Classroom, fostering a dynamic learning environment. As seen in Figure 2, the experiment group students not only received feedback from carefully prepared and theme-appropriate prompts but also received feedback from the teacher, and sometimes this feedback procedure was followed up with a brief chat, further enhancing their immersion in the target language. In contrast, the control group participated in teacher-guided speaking

tasks aligned with the school curriculum, allowing for a comparative analysis of the two approaches.

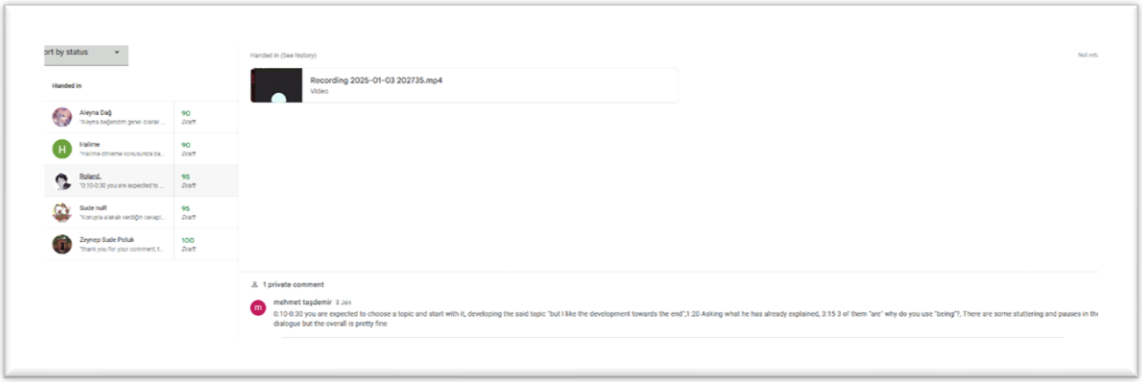


Figure 2 Showing How The Students Uploaded Their Studies And How The Feedback Procedure Went.

The research culminated in the tenth week with the administration of a post-test, where participants once again completed the same instruments used in the pre-test to measure any changes over time. Following the intervention, semi-structured interviews were conducted with members of the experimental group, supplemented by observational notes collected throughout the 10-week process, which provided rich qualitative insights into participant experiences and outcomes.

Data Analysis

The research is a 10-week comprehensive study designed to evaluate the impact of AI-supported speaking activities on participants' FL speaking skills, FL speaking motivation, FL speaking anxiety, and digital literacy. It begins in the first week with a thorough pre-test administration, where all participants undertake a speaking test alongside scales measuring anxiety, motivation, and digital literacy, establishing a solid foundation for subsequent analysis.

For the FL speaking skills, the pre-test and post-test scores were analysed using independent samples t-tests to determine differences between the experimental and control groups, as well as paired samples t-tests to examine within-group changes. For the FL speaking motivation and FL speaking anxiety scores obtained from the validated scales, paired samples t-tests were used to compare pre- and post-intervention results

within groups, and independent samples t-tests were used to identify group-level differences. To further understand the predictive relationships, Pearson correlation analysis was applied between motivation, anxiety, and speaking performance. For digital literacy, independent and paired samples t-tests were conducted similarly, while descriptive statistics were used to provide an overview of participants' competencies. Across all scales, Cronbach's alpha reliability coefficients were calculated to ensure internal consistency.

Following this initial phase, the intervention spans Weeks 2 through 9, during which the experimental group actively engages in innovative, AI-driven speaking exercises facilitated through advanced platforms, such as ChatGPT and Google Gemini. These activities encourage interactive learning, with feedback and discussions managed through Google Classroom to enhance engagement and understanding. Meanwhile, the control group remains focused on traditional, teacher-guided speaking tasks rooted in the existing school curriculum, which allows for a balanced comparison of learning outcomes.

The study culminates in Week 10 with a post-test administration, mirroring the pre-test to assess any shifts in participants' abilities and perceptions over the research period. To enrich the findings further, a series of semi-structured interviews was conducted with members of the experimental group post-intervention, complemented by detailed observational notes collected throughout the 10 weeks. These qualitative insights were analysed thematically and triangulated with quantitative findings, providing a nuanced understanding of the learning experiences and their implications for future educational practices.

To ensure reliability of the findings, triangulation methods were used and data were compared to different sources, these were; interview transcriptions, observation notes and student reflections. Also member checking was used by sharing said themes with participants to verify the accuracy and credibility of the derived themes.

RESULTS

Quantitative Results

Descriptive Statistics

Descriptive analyses were conducted to examine the general trends in students' scores across all measured variables—speaking performance, speaking anxiety, motivation, and digital literacy—prior to and following the AI-supported intervention. The results for all 60 participants are presented in Table 19.

The mean score for speaking performance increased from M = 44.73, SD = 14.59 in the pre-test to M = 54.62, SD = 15.14 in the post-test, indicating an observable improvement in oral proficiency. Similarly, the mean score for motivation rose from M = 253.73, SD = 28.97 to M = 264.95, SD = 28.18, reflecting a positive change in students' language learning motivation over the course of the study.

In contrast, speaking anxiety levels decreased as expected. The pre-test mean for the FLSAS was M = 109.47, SD = 14.66, while the post-test mean dropped to M = 103.35, SD = 17.11, suggesting a reduction in students' anxiety related to speaking English.

Likewise, digital literacy scores demonstrated a meaningful increase from the pre-test (M = 59.25, SD = 8.71) to the post-test (M = 63.92, SD = 9.17), indicating a general enhancement in students' self-reported ability to use digital tools for language learning effectively.

These descriptive findings suggest potential positive effects of AI-supported chatbot interaction on multiple dimensions of students' language learning experience, warranting further analysis through inferential statistics to determine the statistical significance of these observed differences.

Table 19. Descriptive Statistic of Quantitative Results

	N	Min	Max	M	Sd
Pre-Test Digital Literacy	60	34.00	74.00	59.2500	8.70539
Post-Test Digital Literacy	60	41.00	80.00	63.9167	9.17308
Pre-Test FLSAS	60	82.00	136.00	109.4667	14.66168
Post-Test FLSAS	60	74.00	133.00	103.3500	17.11162
Pre-Test Motivation Scale	60	196.00	301.00	253.7333	28.97390
Post-Test Motivation Scale	60	202.00	305.00	264.9500	28.18428
Pre-Test Speaking Exam	60	24.00	75.00	44.7333	14.59495
Post- Test Speaking Exam	60	30.00	86.00	54.6167	15.13901

Descriptive Statistics by Group

A comparison of descriptive statistics for the experimental and control groups (see Tables 20 and 21) reveals clear distinctions in student outcomes following the intervention. As shown in Table 20, students in the experimental group—who engaged in AI-supported speaking activities using tools like ChatGPT and Google Gemini—demonstrated consistent improvements across all measured domains. Their digital literacy scores increased from a pre-test mean of 58.73 to a post-test mean of 68.53, indicating a notable enhancement in their ability to interact with digital tools. Similarly, their Foreign Language Speaking Anxiety Scale (FLSAS) scores decreased from 109.53 to 99.40, reflecting a reduction in speaking anxiety. Motivation levels also rose significantly, with mean scores shifting from 252.93 to 275.20, while speaking performance saw a substantial improvement, rising from 45.70 to 58.43.

Table 20. Descriptive Statistics of Experiment Group

	N	Minimum	Maximum	Mean	Std. Deviation
Pre-Test Digital Literacy	30	34.00	74.00	58.7333	9.37323
Post-Test Digital Literacy	30	55.00	80.00	68.5333	7.64169
Pre-Test FLSAS	30	82.00	131.00	109.5333	15.07163
Pos-Test FLSAS	30	74.00	133.00	99.4000	18.43273
Pre-Test Motivation Scale	30	196.00	301.00	252.9333	29.27036
Post-Test Motivation Scale	30	241.00	297.00	275.2000	18.71087
Pre-Test Speaking Exam	30	25.00	75.00	45.7000	15.48559
Post-Test Speaking Exam	30	35.00	86.00	58.4333	15.39073

In contrast, Table 21 illustrates that the control group, which received traditional instruction without AI integration, experienced only minor changes. Their digital literacy remained nearly unchanged (from 59.77 to 59.30), and speaking anxiety decreased only slightly (from 109.40 to 107.30). Motivation scores stayed almost stable (from 254.53 to 254.70), and although speaking performance improved modestly (from 43.77 to 50.80), the gains were less pronounced than those in the experimental group.

Table 21. Descriptive Statistics of Control Group

	N	Minimum	Maximum	Mean	Std. Deviation
Pre-Test Digital Literacy	30	43.00	73.00	59.7667	8.10995
Post-Test Digital Literacy	30	41.00	74.00	59.3000	8.28855
Pre-Test FLSAS	30	82.00	136.00	109.4000	14.49756
Pos-Test FLSAS	30	81.00	129.00	107.3000	14.95545
Pre-Test Motivation Scale	30	203.00	297.00	254.5333	29.15208
Post-Test Motivation Scale	30	202.00	305.00	254.7000	32.38257
Pre-Test Speaking Exam	30	24.00	72.00	43.7667	13.84317
Post-Test Speaking Exam	30	30.00	71.00	50.8000	14.11627

Together, these descriptive findings provide preliminary evidence supporting the effectiveness of AI-enhanced instruction in promoting not only language performance but also greater digital competence, reduced anxiety, and increased learner motivation. The marked differences observed in Tables 20 and 21 underscore the potential of AI tools to enhance learner engagement and achievement in EFL classrooms compared to conventional pedagogical methods.

Digital Literacy

Before conducting inferential analyses on the digital literacy scores, preliminary tests were carried out to examine the distributional characteristics of the data. Descriptive statistics as seen on Table 22 showed that the mean score for the combined digital literacy scale was 123.17 (SD = 15.42), with a minimum score of 94 and a maximum of 151. The distribution exhibited near symmetry, with a skewness value of .006 (SE = .309), suggesting that the data were not significantly skewed. Kurtosis was –.967 (SE = .608), indicating a relatively flat distribution compared to a normal curve, but still within

acceptable limits for parametric testing. These indicators suggest a reasonably normal distribution of scores across participants.

Table 22. Descriptives of Digital Literacy Normality

		Statistic	Std. Error
Mean		123.1667	1.99125
95% Confidence Interval for Mean	Lower Bound	119.1822	
	Upper Bound	127.1511	
5% Trimmed Mean		123.2407	
Median		122.5000	
Variance		237.904	
Std. Deviation		15.42414	
Minimum		94.00	
Maximum		151.00	
Range		57.00	
Interquartile Range		23.75	
Skewness		.006	.309
Kurtosis		-.967	.608

In table 23 to statistically confirm the assumption of normality, both the Kolmogorov–Smirnov and Shapiro–Wilk tests were performed. The results of the Kolmogorov–Smirnov test yielded a statistic of .068 with a non-significant *p* value of .200, and the Shapiro–Wilk test produced a statistic of .971, $p = .155$. Since both significance values exceed the conventional alpha level of .05, the null hypothesis of normality was retained. As seen on Figure 3 these findings indicate that the digital literacy scores are normally distributed, thereby justifying the use of parametric tests such as the paired samples and independent samples t-tests in subsequent analyses.

Table 23. Tests of Normality for Digital Literacy

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Digital Literacy	.068	60	.200*	.971	60	.155

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

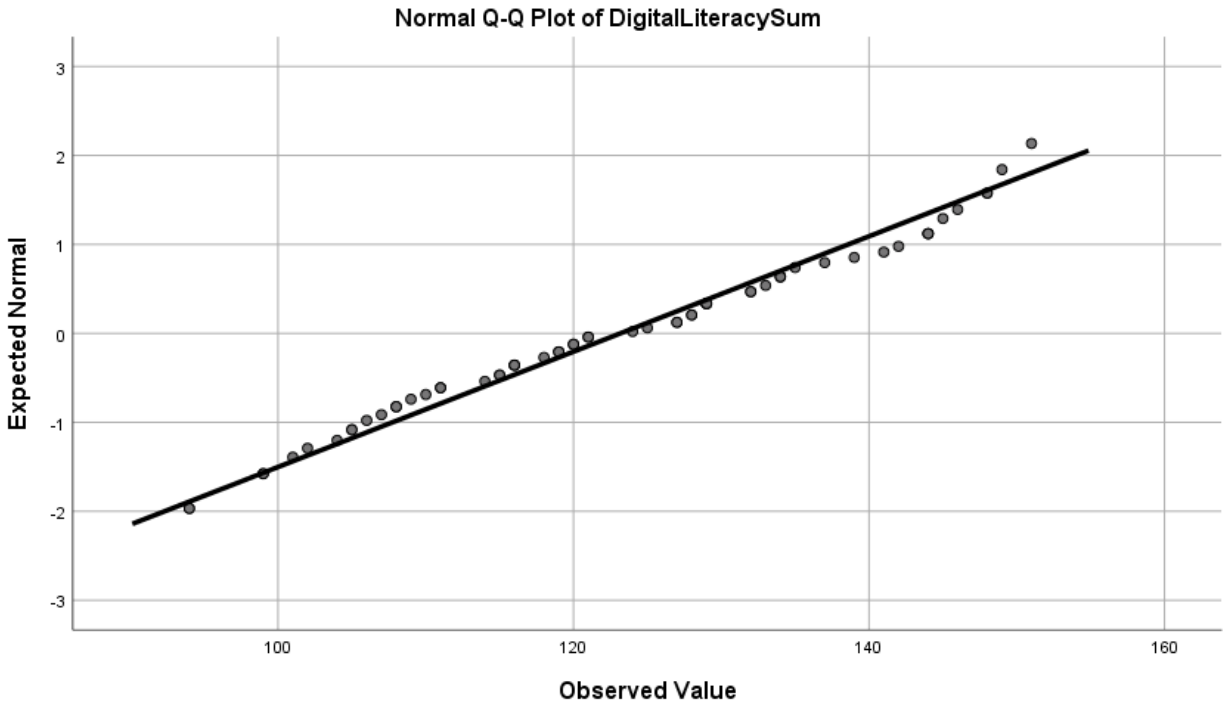


Figure 3 Normality Plot For Digital Literacy

The paired samples t-test of Digital Literacy

The results of the paired samples t-test for digital literacy, as presented in Tables 24–26, demonstrate a significant difference between the experimental and control groups. For the experimental group, Table 24 shows that the mean digital literacy score improved from 58.73 (SD = 9.37) in the pre-test to 68.53 (SD = 7.64) in the post-test. This gain of 9.80 points was statistically significant ($t(29) = –7.73$, $p < .001$), as detailed in Table 26, with a 95% confidence interval ranging from –12.39 to –7.21. The associated effect size (Cohen's $d \approx 1.41$) suggests a very large practical impact, indicating that regular

interaction with AI-based speaking tools like ChatGPT and Google Gemini notably enhanced students' digital literacy over the ten-week period.

Table 24. Results of Digital Literacy Paired Samples Statistics

Group Type		Mean	N	Std. Deviation	Std. Error Mean
Experiment	Pre-Test Digital Literacy	58.7333	30	9.37323	1.71131
	Post-Test Digital Literacy	68.5333	30	7.64169	1.39518
Control	Pre-Test Digital Literacy	59.7667	30	8.10995	1.48067
	Post-Test Digital Literacy	59.3000	30	8.28855	1.51327

Furthermore, Table 25 indicates a strong positive correlation ($r = .684$, $p < .001$) between the pre- and post-test scores in the experimental group, suggesting a consistent pattern of improvement across participants.

Table 25. Results of Digital Literacy Paired Samples Correlations

Group Type		N	Correlation	Sig.
Experiment	Pre-Test Digital Literacy & Post-Test Digital Literacy	30	.684	.000
Control	Pre-Test Digital Literacy & Post-Test Digital Literacy	30	.524	.003

In contrast, the control group exhibited no statistically significant change. As shown in Table 24, their mean scores declined slightly from 59.77 (SD = 8.11) to 59.30 (SD = 8.29), yielding a trivial mean difference of 0.47 points. According to Table 26, this change was not statistically significant ($t(29) = 0.32$, $p = .752$), with the confidence interval (–2.52 to 3.45) indicating that the result could be due to random variation. Although the correlation between scores was moderate ($r = .524$, $p = .003$, Table 25), the effect size was negligible ($d \approx 0.06$), implying that traditional instruction did not impact students' digital literacy in any meaningful way.

Table 26. Results of Digital Literacy Paired Samples Test

Group Type	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Experiment	-9.80000	6.94510	1.26800	-12.39334	-7.20666	-7.729	29	.000
Control	.46667	7.99885	1.46038	-2.52015	3.45349	.320	29	.752

Taken together, these results—supported by data in Tables 24 through 26—highlight the effectiveness of AI-enhanced speaking activities in promoting measurable improvements in digital literacy skills among EFL learners.

Independent Sample T-Tests of Difference in Digital Literacy Skills Between Groups

Pre-test Results

An independent samples t-test was conducted to examine whether there was a significant difference in digital literacy levels between the experimental and control groups prior to the intervention. As shown in Table 27, Levene's test for equality of variances confirmed that the assumption of homogeneity was met ($F(1, 58) = 0.201$, $p = .656$). The mean digital literacy score for the experimental group was 58.73 ($SD = 9.37$), and for the control group, it was 59.77 ($SD = 8.11$). The difference between these means was not statistically significant ($t(58) = –0.46$, $p = .650$), with a 95% confidence interval ranging from –5.56 to 3.50. The calculated Cohen's d of –0.12 indicates a negligible effect size, suggesting no meaningful difference in digital literacy levels between the two groups before the intervention. This baseline equivalence supports the validity of interpreting post-test differences as resulting from the experimental treatment rather than pre-existing disparities.

Table 27. Results of Pre-Test Digital Literacy Independent Samples Test

	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	.201	.656	-.457	58	.650	-1.03333	2.26295	-5.56313	3.49646
Equal variances not assumed			-.457	56.826	.650	-1.03333	2.26295	-5.56512	3.49845

Post-test Results

Following the intervention, another independent samples t-test was conducted to compare post-test digital literacy scores, the results of which are reported in Table 28. Again, Levene's test showed no violation of the equal variance assumption ($F(1, 58) = 0.111$, $p = .740$). The experimental group, having participated in AI-supported speaking activities, achieved a significantly higher mean score ($M = 68.53$, $SD = 7.64$) than the control group ($M = 59.30$, $SD = 8.29$). This difference of 9.23 points was statistically significant ($t(58) = 4.49$, $p < .001$), with a 95% confidence interval from 5.11 to 13.35, indicating that the true mean difference is unlikely due to chance. The effect size was substantial (Cohen's $d = 1.16$), reflecting a large and practically meaningful impact of the AI-based intervention on students' digital literacy development.

Table 28. Results of Post-Test Digital Literacy Independent Samples Test

	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	.111	.740	4.486	58	.000	9.23333	2.05828	5.11324	13.35343
Equal variances not assumed			4.486	57.621	.000	9.23333	2.05828	5.11266	13.35400

These results, documented in Tables 27 and 28, provide strong statistical support for the conclusion that AI-supported speaking activities significantly enhanced digital literacy skills among students in the experimental group, while the control group showed no such gains.

Item-Level Analysis of the Digital Literacy Scale

To assess the impact of the intervention on students' digital literacy which can be seen on Appendix 1, item-by-item changes were analyzed using the data presented in Appendix 8 (Pre-Test) and Appendix 9 (Post-Test). The scale consists of 17 items evaluated on a 5-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree), with higher scores indicating stronger digital literacy perceptions. The analysis revealed that 16 of the 17 items showed increased mean scores, while 1 item showed a decrease, suggesting that students' self-perceived digital literacy generally improved over the course of the intervention.

In the item-level analysis of the Digital Literacy Scale, the categories of notable, moderate, and minimal were determined by considering both statistical and conceptual criteria. Items with high factor loadings (≥ .60), strong item–total correlations, or substantial differences in mean scores were classified as notable, since they contributed significantly to the construct and strongly represented critical dimensions of digital literacy. Items with factor loadings in the range of .40–.59 or moderate correlations were labeled as moderate, reflecting a balanced but less prominent role in defining the scale.

Finally, items with loadings below .40 or weak correlations were regarded as minimal, as their contribution to the construct was relatively limited. Beyond statistical thresholds, content relevance was also taken into account, ensuring that the classification aligned with both psychometric strength and theoretical significance.

Notable Increases

The most substantial increase was observed in Item 7 ("I know how to solve technical problems I encounter"), which rose from M = 3.33 to M = 3.62 (Δ = +0.29), reflecting enhanced confidence in resolving technological difficulties.

Item 9 ("I can keep up with important new technologies") showed an increase of +0.27 points (M = 3.28 to M = 3.55), indicating improved self-efficacy in adapting to technological advancements.

Items 14 and 17, concerning teacher use of technology and knowledge of internet-related topics, each increased by +0.25 points, suggesting enhanced awareness of digital ethics and classroom expectations.

Item 10 ("I have knowledge about many different technologies") rose by +0.23 points.

Item 16 ("My information and communication technology skills are good") increased by +0.20 points.

Item 11 ("I have the technical skills needed to use ICT in learning and creating new things") increased by +0.18 points.

These items indicate significant growth in both technical skills and digital self-efficacy.

Moderate Gains

Several items, including Items 1, 4, 5, 6, 8, and 12, showed modest increases (between +0.08 and +0.15), primarily related to motivation, independent learning, and research abilities in digital environments. This pattern suggests a growing comfort with using digital tools in academic contexts.

Minimal Change or Decline

Item 3 ("Learning using ICT is more engaging") decreased slightly from M = 3.77 to M = 3.50 (Δ = –0.27), which may indicate that initial expectations of engagement were not fully met.

The most significant decrease was seen in Item 13 ("The use of mobile technologies in the learning process has great potential"), which dropped by –0.48 points (M = 3.88 to M = 3.40). This drop might reflect practical limitations experienced by students in mobile learning contexts during the study.

Overall, the results from Appendices 8 and 9 provide compelling evidence of improved digital literacy across nearly all dimensions of the scale. Students reported higher confidence in technical problem-solving, independent digital learning, and adaptation to new technologies. These findings suggest that the intervention (potentially including digital tools or AI-supported components) was effective in fostering digital competence and engagement.

The one exception—Item 13—deserves further investigation, as it may indicate areas where instructional design or infrastructure could be improved to better support mobile learning integration.

FLSAS (Foreign Language Speaking Anxiety Scale) scores.

Before conducting inferential statistical tests, the normality of the total speaking anxiety scores, as measured by the Foreign Language Speaking Anxiety Scale (FLSAS), was assessed using both descriptive and inferential indicators. As presented in Table 29, the data showed approximate symmetry, with a skewness value of 0.017 (SE = 0.309) and a kurtosis of –1.261 (SE = 0.608), indicating no substantial departures from a normal distribution. The mean (M = 212.82) and median (Mdn = 213.00) were nearly identical, further supporting symmetry in the distribution.

Table 29. Descriptives of FLSAS Normality

		Statistic	Std. Error
Mean		212.8167	3.91343
95% Confidence Interval for Mean	Lower Bound	204.9859	
	Upper Bound	220.6474	
5% Trimmed Mean		213.0556	
Median		213.0000	
Variance		918.898	
Std. Deviation		30.31333	
Minimum		158.00	
Maximum		260.00	
Range		102.00	
Interquartile Range		55.75	
Skewness		.017	.309
Kurtosis		-1.261	.608

To statistically assess normality, two tests were applied. The Kolmogorov–Smirnov test returned a non-significant result (D (60) = .114, p = .052), suggesting that the FLSAS scores do not significantly deviate from a normal distribution (see Table 30). Although the Shapiro–Wilk test was statistically significant (W (60) = .943, p = .007), this was interpreted cautiously due to its sensitivity to sample size, particularly in moderate-to-large samples.

Table 30. Tests of Normality for FLSAS

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
FLSAS	.114	60	.052	.943	60	.007

a. Lilliefors Significance Correction

Visual inspection of the distribution, as depicted in Figure 4, supported these findings by illustrating a reasonably linear pattern in the normal Q-Q plot, with no major outliers or systematic deviations. Taken together, the results from Table 29, Table 30, and

Figure 4 suggest that the FLSAS scores are sufficiently normally distributed to justify the use of parametric statistical tests in subsequent analyses.

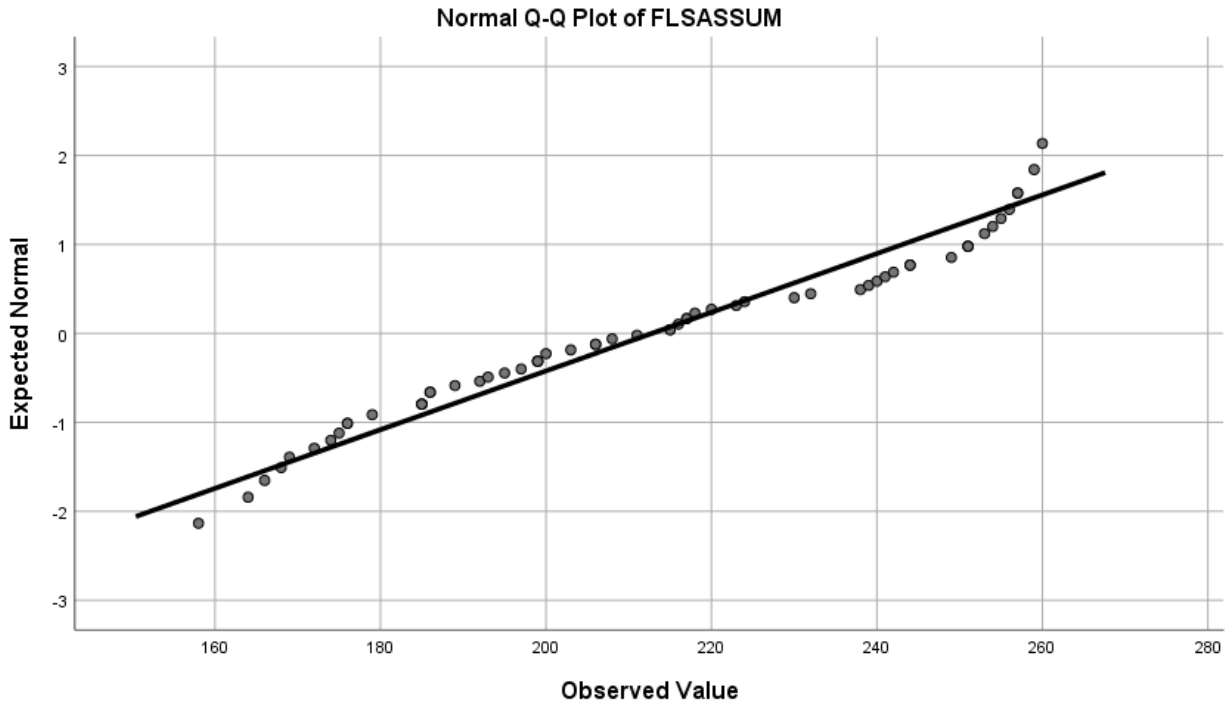


Figure 4 Normality Plot for FLSAS

Paired Samples T-tests

Separate paired samples t-tests were conducted to evaluate changes in foreign language speaking anxiety within the experimental and control groups. As shown in Table 31, the experimental group, which participated in AI-supported speaking activities, experienced a statistically significant reduction in anxiety levels, with mean scores dropping from 109.53 (SD = 15.07) in the pre-test to 99.40 (SD = 18.43) in the post-test. According to Table 33, this mean difference of 10.13 points was significant ($t(29) = 5.66$, $p < .001$), with a 95% confidence interval ranging from 6.47 to 13.80. The corresponding effect size (Cohen's $d = 1.03$) indicates a large and practically meaningful effect, highlighting the efficacy of AI-mediated speaking tasks in reducing learners' speaking anxiety.

Table 31. Paired Samples Statistics of FLSAS

Group Type		Mean	N	Std. Deviation	Std. Error Mean
Experiment	Pre-Test FLSAS	109.5333	30	15.07163	2.75169
	Post-Test FLSAS	99.4000	30	18.43273	3.36534
Control	Pre-Test FLSAS	109.4000	30	14.49756	2.64688
	Post-Test FLSAS	107.3000	30	14.95545	2.73048

Further supporting these results, Table 32 shows a strong correlation between pre- and post-test scores in the experimental group ($r = .847$, $p < .001$), suggesting that improvements were consistent across participants.

Table 32. Paired Samples Correlations of FLSAS

Group Type		N	Correlation	Sig.
Experiment	Pre-Test FLSAS & Post-Test FLSAS	30	.847	.000
Control	Pre-Test FLSAS & Post-Test FLSAS	30	.846	.000

In contrast, the control group showed no statistically significant change in anxiety levels. As presented in Table 31, their pre-test mean was 109.40 (SD = 14.50), and the post-test mean was 107.30 (SD = 14.96). As reported in Table 33, the mean difference of 2.10 points was not significant ($t(29) = 1.41$, $p = .171$), and the 95% confidence interval [–0.96, 5.16] includes zero, suggesting the difference may be due to chance. The effect size was small ($d = 0.26$), indicating only a negligible reduction in anxiety.

Table 33. Paired Samples Tests of FLSAS

Group Type	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Experiment	10.13333	9.81109	1.79125	6.46981	13.79685	5.657	29	.000
Control	2.10000	8.18472	1.49432	-.95622	5.15622	1.405	29	.171

As shown in Table 32, even though the correlation between pre- and post-test scores was also high in the control group ($r = .846$, $p < .001$), the lack of a meaningful difference suggests that traditional instruction did not lead to affective gains. In contrast, the data in Tables 31–33 collectively support the conclusion that AI-supported speaking tasks significantly reduced students' foreign language speaking anxiety.

Independent Sample t-tests

An independent samples t-test was conducted to determine whether there was a significant difference in foreign language speaking anxiety between the experimental and control groups prior to the intervention. As shown in Table 34, Levene's test for equality of variances indicated that the assumption was met ($F(1, 58) = 0.091$, $p = .765$). The mean FLSAS score was 109.53 (SD = 15.07) for the experimental group and 109.40 (SD = 14.50) for the control group. The mean difference of 0.13 points was not statistically significant ($t(58) = 0.035$, $p = .972$), and the 95% confidence interval [–7.51, 7.78] included zero, indicating no meaningful difference. The effect size was minimal (Cohen's $d = 0.01$), confirming that the groups were equivalent in terms of speaking anxiety before the intervention. This statistical equivalence strengthens the internal validity of post-test comparisons by ensuring that any differences observed can be attributed to the intervention rather than baseline disparities.

Table 34. Results of Pre-Test FLSAS Independent Samples Test

	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	.091	.765	.035	58	.972	.13333	3.81809	-7.50940	7.77606
Equal variances not assumed			.035	57.913	.972	.13333	3.81809	-7.50964	7.77631

Following the intervention, another independent samples t-test was conducted to compare the post-test speaking anxiety scores, with results presented in Table 35. Again, Levene's test confirmed homogeneity of variances ($F(1, 58) = 2.540$, $p = .116$). The experimental group, which engaged in AI-supported speaking practice, had a lower mean score ($M = 99.40$, $SD = 18.43$) than the control group ($M = 107.30$, $SD = 14.96$). Although the mean difference of –7.90 points favored the experimental group, the difference did not reach statistical significance ($t(58) = –1.823$, $p = .073$), with a 95% confidence interval ranging from –16.57 to 0.77.

Table 35. Results of Post-Test FLSAS Independent Samples Test

	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	2.540	.116	-1.823	58	.073	-7.90000	4.33371	-16.57486	.77486
Equal variances not assumed			-1.823	55.638	.074	-7.90000	4.33371	-16.58271	.78271

However, the effect size (Cohen's $d = -0.47$) indicates a moderate practical effect, suggesting that AI-supported speaking tasks may have contributed meaningfully to anxiety reduction, even if the result did not achieve conventional levels of statistical significance. The data in Tables 34 and 35 collectively underscore the potential of AI-enhanced instruction to impact learners' affective outcomes, particularly in terms of reducing speaking anxiety.

Item-Level Analysis of the Foreign Language Speaking Anxiety Scale (FLSAS)

To evaluate changes in students' foreign language speaking anxiety, an item-by-item analysis of the FLSAS can bee seen on Appendix 3 was conducted using data from Appendix 10 (Pre-Test Results) and Appendix 11 (Post-Test Results). The scale consisted of 31 items, each rated on a 5-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree), with higher scores generally indicating greater anxiety.

Of the 31 items, 28 exhibited a reduction in mean scores from pre- to post-test, 2 items showed slight increases, and 1 item remained unchanged, suggesting a general decrease in students' speaking anxiety following the intervention. As noted earlier in the item-level analysis of the Digital Literacy Scale, items were categorized as notable, moderate, or minimal based on their factor loadings, item–total correlations, and content relevance. While notable items showed strong statistical and conceptual contributions, moderate items reflected a balanced role, and minimal items had only limited impact on the construct.

Most Substantial Reductions

Item 19 ("I feel my heart pounding when I'm about to speak") decreased from $M = 3.90$ to $M = 3.38$ ($\Delta = -0.52$), indicating a notable reduction in physiological anxiety (Appendices 10–11).

Item 29 ("I'm afraid others will make fun of my speaking") and Item 26 ("I feel very relaxed and confident before class") each declined by 0.43 points, reflecting reduced fear of peer ridicule and lowered anticipatory stress.

Item 3 (“I start trembling when I know I will have to speak”) dropped from M = 3.57 to M = 3.15 (Δ = –0.42), suggesting a marked decrease in somatic symptoms of anxiety.

Item 22 (“I feel very anxious when speaking in front of other students”) showed a 0.38-point reduction, suggesting increased social confidence in classroom interactions.

Moderate Reductions Across Dimensions

Moderate declines were observed in several key items:

Item 13 (“I don’t feel nervous when talking to native speakers”) dropped by 0.34 points, indicating improved comfort in high-pressure communicative situations.

Performance-related fears decreased as well, including Item 9 (“I worry about the consequences of failure”) and Item 8 (“I panic when I have to speak without preparation”), which decreased by 0.28 and 0.36 points, respectively.

Teacher-centered concerns also declined, such as Item 14 (“I get upset when I don’t understand the teacher’s corrections”) and Item 18 (“I think my teacher is ready to correct every mistake”), both of which showed moderate reductions (Δ = –0.12 and –0.26, respectively).

Minimal Change or Slight Increases

Item 21 (“I always feel others speak better than me”) remained unchanged (M = 3.47 pre- and post-test), suggesting ongoing social comparison despite gains in other areas.

Slight increases were observed in Item 11 (“I get so nervous I forget things I know”) and Item 6 (“I sometimes think about unrelated things in class”), though the changes (Δ = +0.02 and +0.03) were negligible and may reflect individual variability.

Overall, the data from Appendices 10 and 11 indicate that the intervention—presumably involving structured speaking practice, possibly supported by AI tools—was associated with meaningful reductions in learners' foreign language speaking anxiety. The most pronounced effects were observed in items reflecting physiological arousal, fear of negative evaluation, and lack of preparedness, aligning with common dimensions of

language anxiety identified in the literature (Horwitz et al., 1986). These findings support the effectiveness of targeted interventions in reducing speaking-related anxiety and fostering a more confident learning environment for EFL learners.

Foreign Language Learning Motivation

Before conducting inferential analyses, the distribution of students' total motivation scores was examined to evaluate the assumption of normality. As displayed in Table 36, the mean motivation score was 518.68 (SD = 38.04), and the distribution showed minimal skewness (0.105, SE = 0.309) and modest negative kurtosis (–0.991, SE = 0.608), indicating an approximately symmetric and slightly platykurtic (flatter-than-normal) distribution. The proximity of the mean (518.68) and median (520.50), along with a relatively modest range (138.00) and interquartile range (61.50), further supports the assumption of a reasonably normal distribution.

Table 36. Descriptives of Foreign Language Learning Motivation Results

		Statistic	Std. Error
Mean		518.6833	4.91041
95% Confidence Interval for Mean	Lower Bound	508.8576	
	Upper Bound	528.5090	
5% Trimmed Mean		518.3333	
Median		520.5000	
Variance		1446.729	
Std. Deviation		38.03588	
Minimum		452.00	
Maximum		590.00	
Range		138.00	
Interquartile Range		61.50	
Skewness		.105	.309
Kurtosis		-.991	.608

The Kolmogorov–Smirnov test, reported in Table 37, was not statistically significant (D(60) = .100, p = .200), suggesting that the distribution did not deviate

significantly from normality. While the Shapiro–Wilk test was statistically significant (W(60) = .959, p = .044), this result was interpreted with caution due to the known sensitivity of the test to sample size in larger datasets.

Table 37. Tests of Normality for Foreign Language Learning Motivation Results

Kolmogorov-Smirnov[a]			Shapiro-Wilk		
Statistic	df	Sig.	Statistic	df	Sig.
.100	60	.200*	.959	60	.044

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

A visual inspection of the data using the normal Q–Q plot (see Figure 5) supported these conclusions, with most data points aligning closely along the diagonal reference line and no clear outliers or systematic deviations.

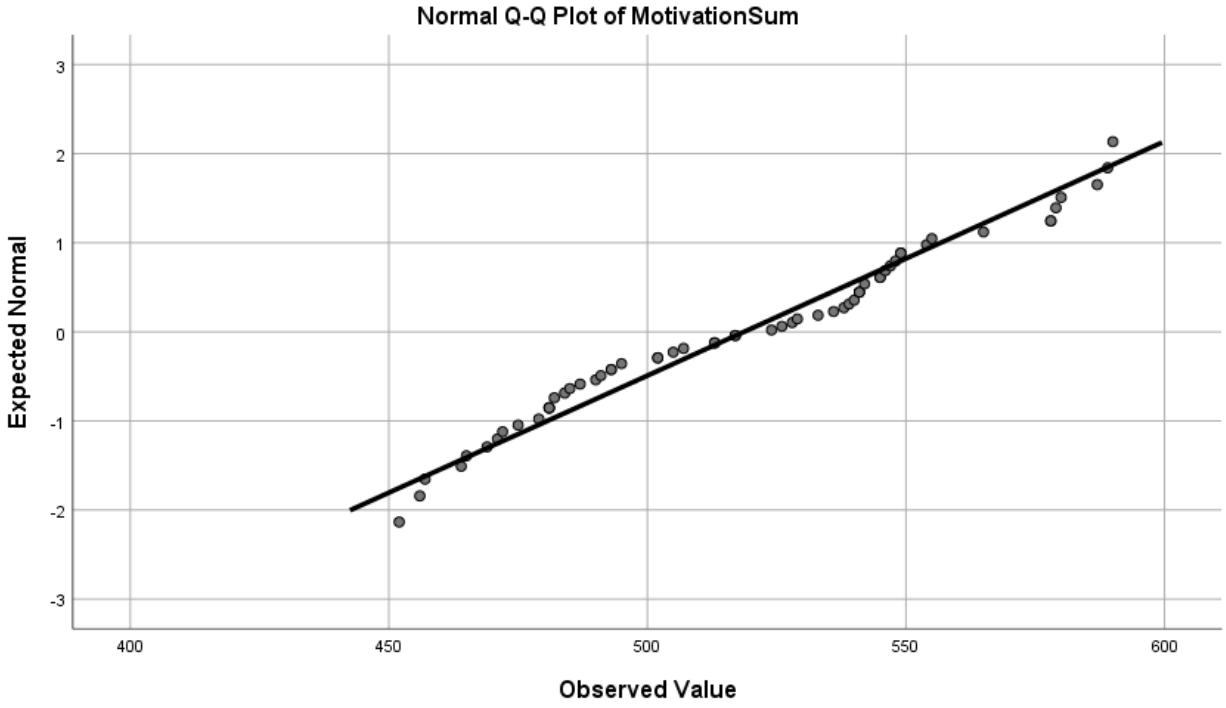


Figure 5 Normality Plot for Foreign Language Leaning Motivation

Taken together, the statistical results in Tables 36 and 37, along with the graphical evidence from Figure 5, suggest that the motivation scores are sufficiently normally distributed to justify the use of parametric statistical tests in subsequent analyses.

Paired Sample T-Test

Paired samples t-tests were conducted to assess changes in language learning motivation within the experimental and control groups. As shown in Table 38, the experimental group, which received AI-supported speaking instruction, exhibited a substantial increase in motivation scores, rising from a pre-test mean of 252.93 (SD = 29.27) to a post-test mean of 275.20 (SD = 18.71). According to Table 40, this mean difference of –22.27 points was statistically significant, $t(29) = –6.19$, $p < .001$, with a 95% confidence interval ranging from –29.62 to –14.91. The corresponding effect size (Cohen's d = 1.13) indicates a large and practically meaningful improvement in motivation following the AI-enhanced intervention.

Table 38. Paired Samples Statistics of Foreign Language Learning Motivation Scale

Group Type		Mean	N	Std. Deviation	Std. Error Mean
Experiment	Pre-Test Motivation Scale	252.9333	30	29.27036	5.34401
	Post-Test Motivation Scale	275.2000	30	18.71087	3.41612
Control	Pre-Test Motivation Scale	254.5333	30	29.15208	5.32242
	Post-Test Motivation Scale	254.7000	30	32.38257	5.91222

Additionally, Table 39 shows a strong positive correlation ($r = .748$, $p < .001$) between the pre- and post-test motivation scores, supporting the consistency of the observed increase across participants.

Table 39. Paired Samples Correlations of Foreign Language Learning Motivation Scale

Group Type		N	Correlation	Sig.
Experiment	Pre-Test Motivation Scale & Post-Test Motivation Scale	30	.748	.000
Control	Pre-Test Motivation Scale & Post-Test Motivation Scale	30	-.617	.000

In contrast, the control group displayed no significant change in motivation over the same period. As reported in Table 38, their mean motivation score remained virtually unchanged—from 254.53 (SD = 29.15) at pre-test to 254.70 (SD = 32.38) at post-test. Table 40 indicates that this negligible difference (–0.17 points) was not statistically significant ($t(29) = –0.016$, $p = .987$), with a 95% confidence interval spanning –20.84 to 20.50, and a trivial effect size ($d = 0.003$). Interestingly, Table 39 reveals a moderate negative correlation ($r = –.617$, $p < .001$) between pre- and post-test scores in the control group, which may reflect variability unrelated to the instructional context.

Table 40. Paired Samples Test Results of Foreign Language Learning Motivation Scale

Group Type	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Experiment	-22.26667	19.69235	3.59531	-29.61991	-14.91342	-6.193	29	.000
Control	-.16667	55.35595	10.10657	-20.83692	20.50359	-.016	29	.987

Overall, the findings reported in Tables 38–40 strongly suggest that participation in AI-supported speaking tasks significantly enhanced learners' motivation to study English, whereas traditional instruction produced no measurable motivational gains.

Independent Samples T-test

An independent samples t-test was conducted to assess whether there were any baseline differences in language learning motivation between the experimental and control groups prior to the intervention. Levene's test for equality of variances was not significant, $F(1, 58) = 0.133$, $p = .717$, indicating that the assumption of homogeneity of variances was met. The experimental group had a mean pre-test motivation score of $M = 252.93$, $SD = 29.27$, while the control group had a mean of $M = 254.53$, $SD = 29.15$ (see Table 41).

Table 41. Group Statistics of Pre-Test Motivation Scale

GroupType	N	Mean	Std. Deviation	Std. Error Mean
Experiment	30	252.9333	29.27036	5.34401
Control	30	254.5333	29.15208	5.32242

The t-test revealed no statistically significant difference between the groups, $t(58) = –0.212$, $p = .833$, 95% CI [–16.70, 13.50] (see Table 42). The effect size was negligible, Cohen's $d = –0.05$, suggesting that both groups were equivalent in terms of motivation at the outset of the study.

Table 42. Independent Samples Test for Pre-Test Motivation Scale

	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	.133	.717	-.212	58	.833	-1.60000	7.54232	-16.69759	13.49759
Equal variances not assumed			-.212	57.999	.833	-1.60000	7.54232	-16.69760	13.49760

A second independent samples t-test was carried out to examine differences in post-test motivation scores between the groups. Levene's test indicated a violation of the assumption of equal variances, $F(1, 58) = 8.886$, $p = .004$; therefore, the Welch's t-test was used. The results showed a statistically significant difference in favor of the experimental group, which scored higher ($M = 275.20$, $SD = 18.71$) than the control group ($M = 254.70$, $SD = 32.38$) (see Table 43).

Table 43. Group Statistics of Post-Test Motivation Scale

Group Type	N	Mean	Std. Deviation	Std. Error Mean
Experiment	30	275.2000	18.71087	3.41612
Control	30	254.7000	32.38257	5.91222

The mean difference of 20.50 points was significant, $t(46.42) = 3.00$, $p = .004$, with a 95% confidence interval of [6.76, 34.24] (see Table 44). The calculated Cohen's d was 0.78, indicating a large effect size and a meaningful improvement in language learning motivation attributed to the AI-supported intervention.

Table 44. Independent Samples Test for Post-Test Motivation Scale

	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	8.886	.004	3.002	58	.004	20.50000	6.82820	6.83188	34.16812
Equal variances not assumed			3.002	46.422	.004	20.50000	6.82820	6.75893	34.24107

These findings confirm that the experimental group benefited significantly from the AI-enhanced instruction in terms of motivation, whereas no such improvement was observed at baseline, affirming the validity of the intervention's impact.

Item-Level Analysis of the Foreign Language Motivation Scale

The pre- and post-test results for the Foreign Language Motivation Scale which can be seen on Appendix 2 indicate significant shifts in students' motivational profiles across multiple dimensions. Out of 72 items, 68 demonstrated an increase, 3 showed a decrease, and 1 remained unchanged, indicating an overall enhancement in motivation levels following the intervention.

Notable Increases

The largest increase was seen in Item 48 ("After finishing this course, I will probably take another English course"), rising from M = 3.27 to M = 3.83 (Δ = +0.56). This suggests a meaningful boost in students' long-term motivation and commitment to language learning.

Items 5 and 4, related to intrinsic motivation ("I wish I could learn English more easily without attending class" and "I don’t enjoy learning English but I know it is important"), increased by +0.55 and +0.54 respectively. This points to stronger internalization of language learning goals.

Item 54 ("In English class, the teacher should speak and students should respond only when called upon") increased by +0.50, possibly indicating a growing appreciation for teacher-led structure after the instructional experience.

Similarly, Item 56 ("Students should ask questions whenever they don’t understand") rose by +0.49, reflecting stronger beliefs in learner autonomy and classroom participation.

Item 44 ("I often struggle to stay focused in English class") from 3.45 to 3.82 (Δ = +0.37), indicating improved classroom engagement.

Item 39 ("I feel uncomfortable when I have to speak in English class") actually decreased from 3.65 to 3.40 (Δ = –0.25), reflecting a reduction in speaking anxiety, which aligns with earlier findings from the FLSAS.

Stable or Slight Changes

A number of items, such as Item 2 ("Learning English is a hobby for me") and Item 8 ("I want to succeed in this course to show my skills to others"), showed minimal or no change ($\Delta \leq \pm 0.05$), suggesting these constructs were stable throughout the study.

Item 12 ("I want to learn English because it is useful when traveling") decreased slightly ($\Delta = -0.19$), though the mean remained high ($M = 3.82 \rightarrow 3.63$), indicating sustained instrumental motivation.

Decreases

Item 39 showed the largest decrease (–0.25), which, while a numerical drop, is actually positive, as it reflects reduced classroom anxiety.

Item 38 ("British culture has contributed greatly to the world") dropped slightly (–0.24), which might suggest shifting cultural attitudes rather than motivational decline.

Item 33 ("If I succeed in this class, it will be because the class is easy") also decreased (–0.20), indicating that students may attribute success more to effort or teaching quality rather than course difficulty.

The results presented in Appendices 6 and 7 reveal that the intervention positively influenced a broad range of motivational constructs: intrinsic interest, instrumental goals, expectancy beliefs, and motivational strength. The most substantial improvements appeared in areas related to future intention to study, engagement in class activities, and autonomous learning behaviors. Additionally, subtle reductions in items associated with anxiety and external attributions suggest a more internalized and empowered view of English language learning post-intervention.

These findings support the effectiveness of the learning environment—potentially enhanced through AI-assisted or communicative practices—in not only reducing anxiety but also fostering sustained motivational growth across cognitive, behavioral, and affective dimensions.

Speaking Performance

To determine the appropriateness of parametric statistical analyses, the distributional characteristics of the speaking performance scores were thoroughly evaluated. Descriptive statistics revealed that the mean score was M = 99.35 with a standard deviation of SD = 28.77, and a median value of 92.50, suggesting a moderately centered distribution (see Table 45). The range of scores extended from 57 to 149, with a 5% trimmed mean of 98.85, indicating minimal influence from outliers. The skewness value of 0.300 (SE = 0.309) indicated a slight positive skew, while kurtosis was –1.182 (SE = 0.608), pointing to a flatter-than-normal distribution.

Table 45. Descriptives of Speaking Scores

		Statistic	Std. Error
Mean		99.3500	3.71406
95% Confidence Interval for Mean	Lower Bound	91.9182	
	Upper Bound	106.7818	
5% Trimmed Mean		98.8519	
Median		92.5000	
Variance		827.655	
Std. Deviation		28.76900	
Minimum		57.00	
Maximum		149.00	
Range		92.00	
Interquartile Range		49.00	
Skewness		.300	.309
Kurtosis		-1.182	.608

Although both normality tests were statistically significant—Kolmogorov–Smirnov: D(60) = .127, p = .018, and Shapiro–Wilk: W(60) = .927, p = .001—suggesting deviations from a normal distribution (see Table 46), further inspection of the distribution (see Figure 6) supported the conclusion that the data did not contain severe violations of normality. Importantly, the relatively symmetric shape, lack of substantial skewness or leptokurtosis, and the sample size of N = 60 supported the robustness of using parametric

tests, such as paired and independent samples t-tests, which are generally reliable in samples of this size.

Table 46. Tests of Normality for Speaking Scores

Kolmogorov-Smirnov[a]			Shapiro-Wilk		
Statistic	df	Sig.	Statistic	df	Sig.
.127	60	.018	.927	60	.001

a. Lilliefors Significance Correction

Despite the statistical significance of the normality tests, the combination of visual inspection (see Figure 6), robust sample size, and moderate distributional properties justifies the use of parametric methods in subsequent analyses.

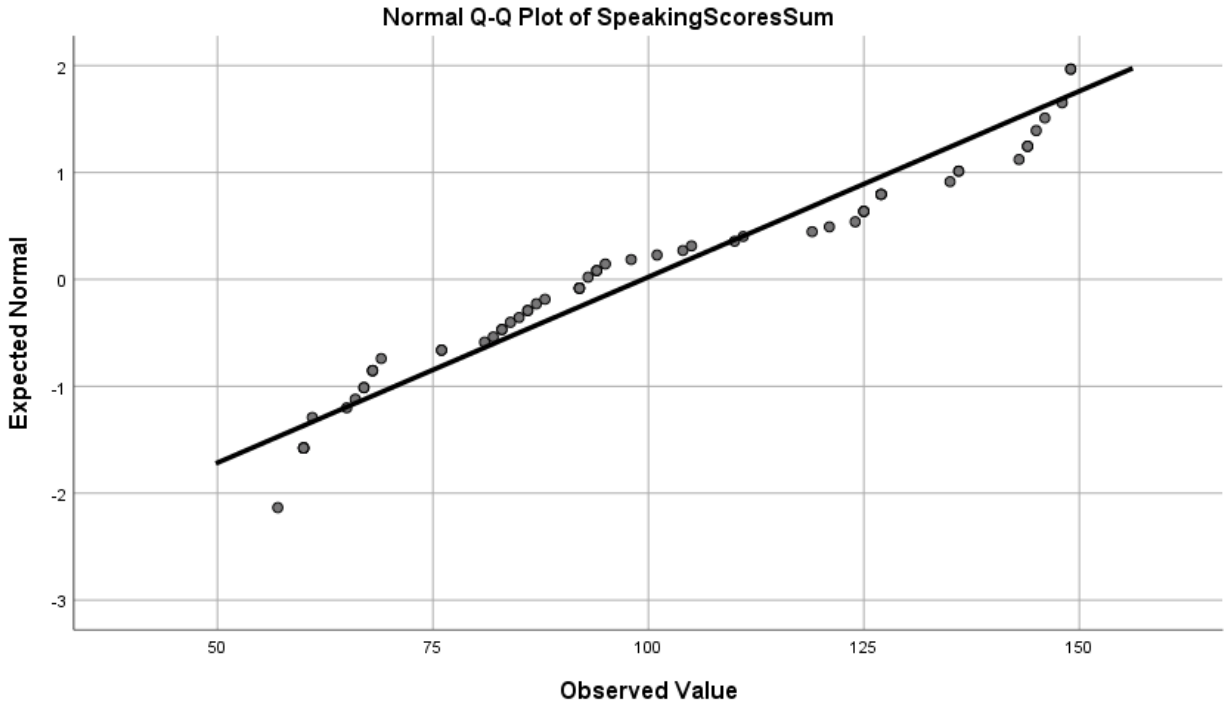


Figure 6 Normality Plot for Speaking Scores

Paired Samples T-Test

Paired samples t-tests were conducted to assess within-group changes in speaking performance from pre-test to post-test for both the experimental and control groups. As shown in Table 47, the experimental group, which received AI-supported speaking

instruction, improved from a mean pre-test score of 45.70 (SD = 15.49) to a mean post-test score of 58.43 (SD = 15.39). This change was statistically significant, $t(29) = –9.29$, $p < .001$, with a 95% confidence interval of [–15.54, –9.93] (see Table 49). The effect size, calculated using Cohen's d, was 1.69, which represents a very large and practically meaningful improvement. Additionally, the paired samples correlation was strong ($r = .882$, $p < .001$), indicating a consistent relationship between pre- and post-test scores within the group (see Table 48).

Table 47. Paired Samples Statistics of Speaking Scores

Group Type	Mean	N	Std. Deviation	Std. Error Mean
Experiment	45.7000	30	15.48559	2.82727
	58.4333	30	15.39073	2.80995
Control	43.7667	30	13.84317	2.52741
	50.8000	30	14.11627	2.57727

Similarly, the control group showed a statistically significant improvement in speaking performance. The mean score increased from 43.77 (SD = 13.84) to 50.80 (SD = 14.12), as shown in Table 47. The paired t-test revealed a significant difference, $t(29) = –5.92$, $p < .001$, with a 95% confidence interval of [–9.46, –4.60] (see Table 49).

Table 48. Paired Samples Correlations of Speaking Scores

Group Type	N	Correlation	Sig.
Experiment	30	.882	.000
Control	30	.892	.000

The effect size for the control group was Cohen's $d = 1.08$, which also indicates a large effect. The paired samples correlation was similarly high, $r = .892$, $p < .001$ (see Table 48), reflecting stable within-group progression.

Table 49. Paired Samples Test of Speaking Scores

Group Type	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Experiment	-12.73333	7.50601	1.37040	-15.53612	-9.93054	-9.292	29	.000
Control	-7.03333	6.50455	1.18756	-9.46217	-4.60449	-5.922	29	.000

Taken together, both groups improved significantly in their speaking scores. However, the experimental group outperformed the control group both in terms of mean gain (ΔM = 12.73 vs. 7.03) and effect size (1.69 vs. 1.08). These results support the conclusion that AI-supported speaking practice had a greater impact on students' speaking performance than traditional methods.

Independent Samples T-test

An independent samples t-test was performed to assess whether any pre-existing differences in speaking performance existed between the experimental and control groups prior to the intervention. As shown in Table 50, the experimental group had a mean pre-test score of 45.70 (SD = 15.49), while the control group had a mean of 43.77 (SD = 13.84). Levene's test for equality of variances was not significant, $F(1, 58) = 1.037$, $p = .313$, indicating that the assumption of equal variances was met.

Table 50. Group Statistics of Pre-Test Speaking Scores

	GroupType	N	Mean	Std. Deviation	Std. Error Mean
SpeakingExamPreTest	Experiment	30	45.7000	15.48559	2.82727
	Control	30	43.7667	13.84317	2.52741

The independent samples t-test revealed no statistically significant difference, $t(58) = 0.51$, $p = .612$, with a 95% confidence interval of [–5.66, 9.52] (see Table 51).

The effect size, measured using Cohen's d, was 0.13, suggesting a small and practically negligible difference. These results confirm that the two groups were comparable in terms of initial speaking performance.

Table 51. Independent Samples Test of Pre-Test Speaking Scores

	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	1.037	.313	.510	58	.612	1.93333	3.79226	-5.65770	9.52436
Equal variances not assumed			.510	57.286	.612	1.93333	3.79226	-5.65971	9.52638

A second independent samples t-test was conducted to compare post-test speaking scores between the experimental and control groups. According to Table 52, the experimental group, which had received AI-supported speaking practice, achieved a mean score of 58.43 (SD = 15.39), whereas the control group scored 50.80 (SD = 14.12). Levene's test confirmed equal variances, $F(1, 58) = 0.395$, $p = .532$, validating the use of the standard t-test.

Table 52. Group Statistics of Post-Test Speaking Scores

Group Type	N	Mean	Std. Deviation	Std. Error Mean
Experiment	30	58.4333	15.39073	2.80995
Control	30	50.8000	14.11627	2.57727

The results showed a statistically significant difference between groups, $t(58) = 2.00$, $p = .050$, with a mean difference of 7.63 points and a 95% confidence interval ranging from 0.001 to 15.27 (see Table 53). The effect size, calculated as Cohen's $d = 0.52$, reflects a moderate and educationally meaningful difference in favor of the

experimental group. This suggests that AI-enhanced speaking instruction provided added value over traditional methods in improving students' speaking performance.

Table 53. Independent Samples Test of Post-Test Speaking Scores

	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	.395	.532	2.002	58	.050	7.63333	3.81289	.00101	15.26566
Equal variances not assumed			2.002	57.572	.050	7.63333	3.81289	-.00020	15.26687

Qualitative Results

In the study, semi-structured interviews were conducted with ten students from the experimental group to gain deeper insights into their experiences with AI-supported speaking practice. Utilizing Braun and Clarke's (2006) thematic analysis method, the research unearthed six key themes that encapsulate the students' perspectives. Firstly, many students reported a significant reduction in speaking anxiety thanks to non-judgmental interactions facilitated by AI, which created a safer environment for practicing their skills. Additionally, the immediacy and accessibility of feedback provided by the AI tools were highlighted as beneficial, allowing students to receive real-time corrections and suggestions that fostered improvement. However, the limitations of AI in contextual understanding and emotional responsiveness were also noted, indicating that while the technology offers support, it may not fully grasp the nuances of human conversation. Furthermore, students expressed a desire for greater learner autonomy and the opportunity to engage in increased speaking practice, which the AI tools helped facilitate. Many participants voiced their preference for hybrid and personalized learning models that could combine traditional and AI-assisted methods, aiming for a more tailored educational experience. Lastly, some digital challenges and linguistic limitations

surfaced, suggesting areas for further development in AI technology within language learning contexts.

Reduction of Speaking Anxiety through Non-Judgmental Interaction

A dominant theme across interviews was the students' sense of emotional comfort when interacting with AI chatbots. Many participants emphasized that speaking with AI felt less intimidating than speaking with teachers or peers, reducing their anxiety and encouraging more frequent participation. One of the underlying reasons was the absence of direct human judgment. Learners often associated classroom speaking activities with the possibility of being corrected in front of others, which created a sense of exposure and vulnerability. In contrast, when interacting with an AI chatbot, they reported feeling less pressure, since the interaction was perceived as private, non-judgmental, and forgiving of mistakes. Several students explained that while they feared being laughed at or misunderstood by their peers, AI did not "laugh," "show emotions," or "embarrass" them, which made the process safer.

> *"It was really fun! I wasn't afraid of making mistakes because the AI didn't feel like it was judging me. I could try again and again." (Learner 1)*
>
> *"Since I wasn't scared of making mistakes with the AI, I practiced more. This gave me more courage to speak in class too." (Learner 9)*
>
> *"After practicing with ChatGPT at home, I started to feel less nervous even when speaking in class." (Learner 4)*

This reflects Krashen's (1982) affective filter hypothesis, suggesting that anxiety-free environments facilitate language acquisition by lowering emotional barriers.

Immediacy and Accessibility of Feedback

Students frequently mentioned the advantage of receiving instant feedback from AI. Unlike classroom settings where feedback might be delayed due to time constraints, AI provided immediate corrections and responses, which was perceived as beneficial. Many participants explained that this immediacy allowed them to identify their mistakes while the context was still fresh in their minds, making it easier to understand and remember the correction. They emphasized that in regular classrooms, the teacher's

attention is divided among many students, and feedback often comes after a noticeable delay, sometimes at the end of an activity or even the following day. This delay, they noted, lessened the impact of the feedback and made it harder to connect corrections with their actual performance.

> *"What I liked the most was getting instant feedback! With teachers, you have to wait your turn, but AI responds immediately." (Learner 9)*

> *"It was really convenient to just type something into ChatGPT and get a reply. It even helped me with homework." (Learner 4)*

This theme supports the pedagogical value of *real-time feedback*, known to enhance retention and engagement (Shute, 2008).

Limitations of AI in Contextual Understanding and Emotional Responsiveness

While students valued the speed and consistency of AI feedback, several expressed dissatisfactions with its depth and human-like understanding. Some found the responses robotic or lacking natural flow, and others noted that AI primarily focused on grammar without contextual explanations. For instance, a number of learners described how AI corrections sometimes felt mechanical—simply pointing out an error and providing the "right" form, without offering insight into why the mistake occurred or how to avoid it in future contexts. They contrasted this with human teachers, who often connected corrections to real-life situations, cultural nuances, or pragmatic aspects of language use, making the learning experience more meaningful.

> *"AI only corrects grammar. My teacher explains why a sentence is incorrect and how to make it more natural. AI doesn't do that." (Learner 8)*

> *"Sometimes the answers from AI were strange. Talking to a real person feels more natural." (Learner 5)*

> *"It would be more motivating if the AI could say things like 'Well done!' It's too robotic right now." (Learner 10)*

These limitations highlight the current gap between AI's linguistic accuracy and its pragmatic or affective communication capabilities.

Learner Autonomy and Increased Speaking Practice Opportunities

A number of students reported that AI tools empowered them to practice outside of classroom hours. This independent usage not only increased speaking frequency but also boosted confidence. Learners explained that having 24/7 access to a conversational partner eliminated one of the biggest limitations of traditional classroom learning: restricted practice time. In school, they often had to wait for designated speaking activities, and even then, opportunities were unevenly distributed among classmates. With AI, however, they could engage in speaking practice anytime, anywhere, without needing a teacher or peer present.

> *"I talked to ChatGPT at home too. After so much practice, my confidence improved." (Learner 4)*
>
> *"I practiced more with AI because I didn't feel shy." (Learner 7)*

This finding aligns with the concept of *learner autonomy* (Little, 1995), which emphasizes the benefits of self-directed language learning, especially when supported by digital tools.

Desire for Hybrid and Personalized Learning Models

Several participants suggested that the ideal approach would be a combination of AI tools and teacher support. They believed that AI is well-suited for practice and repetition, while teachers provide deeper, more nuanced instruction. Students frequently explained that although AI gave them unlimited opportunities to rehearse language, it could not always address the complexity of meaning, cultural context, or pragmatic appropriateness. Teachers, on the other hand, were seen as capable of offering richer explanations, personal encouragement, and context-sensitive guidance that AI currently lacks.

> *"Practicing with AI helped improve my fluency, but for pronunciation I still needed my teacher's corrections." (Learner 3)*
>
> *"If we had one session with AI and one with a teacher each week, that would be perfect." (Learner 7)*
>
> *"AI is great for practice; teachers are better for deep learning. Having both would be amazing." (Learner 10)*

These insights support blended learning models (Graham, 2006), which combine traditional and digital approaches to address varied learner needs.

Digital Challenges and Linguistic Limitations

While students generally adapted well to the AI platforms, a few reported technical or linguistic difficulties, particularly in pronunciation recognition and comprehension of accents.

> *"I wish the AI could understand the Turkish accent better. When I said 'apple' like 'epıl', it got confused." (Learner 6)*

> *"Sometimes the AI didn't understand me and I had to repeat myself again and again. It was frustrating." (Learner 2)*

These challenges point to the importance of *inclusive and adaptive design* in educational AI, ensuring that learners with different backgrounds or pronunciation patterns are not disadvantaged.

Table 54. Summary of Qualitative Themes

Theme	Description	Illustrative Quote
1.Reduced Anxiety	Students felt emotionally safe with AI	"AI didn't feel like it was judging me."
2.Instant Feedback	AI gave real-time, motivating responses	"AI responds immediately."
3. Feedback Gaps	AI lacked contextual and emotional depth	"AI is too robotic."
4. Autonomy	AI encouraged out-of-class practice	"I practiced more with AI because I wasn't shy."
5.Hybrid Preference	Combining AI and teacher was ideal	"AI for practice, teacher for deep learning."
6.Technical Issues	Some challenges with accents or comprehension	"AI didn't understand my pronunciation.

DISCUSSION

This study aimed to investigate the effects of AI-supported chatbot interaction on EFL high school students' speaking performance, speaking anxiety, motivation, and digital literacy. The discussion in this section evaluates the study's findings in light of the existing literature and theoretical frameworks while considering the study's limitations.

The quantitative results demonstrated that AI-mediated speaking practice led to statistically significant improvements in speaking performance, motivation, and digital literacy, along with a meaningful decrease in speaking anxiety. These outcomes are further supported by the qualitative findings, which revealed that students perceived AI as a non-judgmental, accessible, and motivating tool for oral language practice. This convergence between statistical analysis and learner perspectives suggests a robust pedagogical value in integrating AI tools into language instruction (Shazly, 2021).

A notable finding was the reduction in students' speaking anxiety after interacting with AI chatbots. This supports previous studies such as (Almineeai et al., 2025), who found that computer-mediated communication helped reduce anxiety among language learners by offering a less judgmental environment. In line with Krashen's Affective Filter Hypothesis (1982), the present study demonstrates that anxiety-reducing environments can facilitate more effective language acquisition. The qualitative data confirmed that learners felt safer making mistakes with AI than with human interlocutors. This aligns with the work of Busso and Sanchez, (2024) who highlighted the importance of emotional safety in encouraging oral participation among anxious learners.

In terms of speaking performance, the observed improvements were consistent with prior findings by Lee et al., (2023), who reported that AI-supported tools helped EFL learners develop fluency and vocabulary use. Similarly, Kemelbekova et al., (2024) noted that chatbot interactions can simulate authentic communication and thereby promote productive language output. In the current study, participants attributed their improved fluency to the repetitive, low-pressure practice made possible by AI, echoing the benefits described in Long's Interaction Hypothesis (1996), where negotiation of meaning during interaction fosters language development. Nevertheless, some students indicated that AI lacked contextual depth in feedback, particularly regarding pronunciation and discourse naturalness—an observation that confirms the limitations identified in Young and

Shishido (2023), who argue that AI lacks pragmatic sensitivity and cannot replace teacher-mediated feedback entirely.

Motivation also emerged as a strengthened domain, particularly among the experimental group, which aligns with the motivational constructs outlined in Self-Determination Theory (Deci & Ryan, 1985). AI tools appeared to support both intrinsic motivation (through novelty, interactivity, and instant feedback) and extrinsic motivation (by helping students see measurable improvement). This is consistent with the findings of Dang (2025), who showed that mobile-based AI interactions increased learners' sense of autonomy and engagement. However, the correlation analysis in this study revealed that motivation was not significantly related to performance or digital literacy in the post-test, suggesting that while motivation increased overall, its impact may have varied individually or been mediated by other affective factors. This could be due to the novelty effect of the technology or individual preferences toward digital learning, which are not always captured in standardized scales.

Digital literacy was another area where significant gains were observed in the experimental group, consistent with previous studies like Hamutoğlu et al. (2017), who emphasized the potential of digital learning tools to enhance learners' confidence and competence in navigating educational technologies. Students' reflections indicated that AI platforms such as ChatGPT and Google Gemini offered frequent, self-paced opportunities for language practice, which indirectly contributed to greater familiarity and comfort with digital environments. However, several students reported challenges in pronunciation recognition and accent comprehension, reinforcing claims by Liu and Darvin (2024) that AI tools still require improvements in accommodating non-native accents and regional speech patterns.

The preference for a hybrid learning model, as voiced by multiple students in the qualitative phase, highlights the nuanced role AI can play in modern classrooms. While students appreciated the accessibility and feedback speed of AI, they did not perceive it as a replacement for teacher guidance. Instead, they emphasized the complementary role of human instruction, especially in providing personalized, empathetic, and contextually rich feedback. This supports Liu and Mantuhac (2024) model of blended learning,

suggesting that the most effective outcomes may arise when AI tools are used not in isolation, but alongside teacher-led instruction.

Despite the promising findings, the study is not without limitations. The relatively short duration (10 weeks) may not have been sufficient to capture long-term retention or sustained affective change. The study was also limited to 11th-grade students in one public high school, which may affect the generalizability of results to other age groups or educational contexts. Additionally, although the instruments demonstrated strong reliability, self-reported data—particularly for constructs like motivation and digital literacy—may be influenced by students' self-perception or social desirability. Furthermore, the study only focused on two specific AI tools; the inclusion of more diverse platforms could have broadened the understanding of AI's capabilities and limitations.

In conclusion, the results of this study align well with a growing body of research supporting the integration of AI technologies into language education. The findings suggest that AI chatbots can effectively support the development of speaking skills, reduce anxiety, and foster motivation and digital competence—particularly when thoughtfully embedded in a hybrid instructional model. Future studies could extend this research by exploring longitudinal outcomes, teacher perceptions, and the impact of AI on other language skills such as listening and writing, as well as its role in assessment.

CONCLUSION & RECOMMENDATIONS

This study examined the effects of AI-supported chatbot interaction on high school EFL learners' speaking performance, speaking anxiety, motivation, and digital literacy. Based on the analysis of both quantitative and qualitative data, it was concluded that students who engaged in AI-supported speaking activities showed notable improvement in their overall speaking skills. Their performance gains were especially evident in fluency and vocabulary use, suggesting that frequent interaction with AI tools contributed positively to oral language development. The results also indicated a significant decrease in students' foreign language speaking anxiety. Learners expressed that speaking with AI tools felt less stressful than speaking with peers or teachers, enabling them to practice more comfortably and with increased confidence.

In addition to language performance, the study showed that students' motivation toward speaking English improved after engaging with AI chatbots. Participants became more willing to participate in speaking tasks, partly due to the instant feedback and continuous practice opportunities that the AI provided. Furthermore, students' digital literacy skills were enhanced, as they became more familiar and confident in using digital platforms and tools during the learning process. The qualitative data supported these outcomes, revealing that most students viewed AI as an effective and supportive tool for language learning. However, they also expressed the need for teacher involvement, especially in providing deeper explanations and personalized feedback, indicating that a balance between AI and teacher support would offer the most comprehensive learning experience.

In light of these findings, several recommendations are offered for researchers and practitioners. For researchers, future studies may explore the long-term effects of AI-supported speaking practice to determine whether the improvements observed in this study are sustainable over time. Research conducted with different age groups, such as younger learners or university students, could help determine whether similar results are found across educational levels. It may also be beneficial to compare different AI tools—such as voice-based and text-based systems—or to evaluate the integration of AI in skills beyond speaking, such as listening or pronunciation. Additionally, further qualitative

studies involving teachers' perceptions could provide a richer understanding of how AI can be most effectively integrated into language instruction.

For practitioners working in EFL classrooms, it is recommended that AI chatbots like ChatGPT and Google Gemini be incorporated as a complementary tool for speaking practice. Teachers should consider using these platforms to provide students with opportunities for autonomous, low-anxiety practice, particularly for learners who may be hesitant to speak in front of others. However, AI tools should not be used in isolation. Instead, a hybrid model in which AI-supported tasks are supplemented by teacher-led discussions and corrective feedback would help address the limitations of AI and maximize the benefits for learners. Teachers are also encouraged to assign AI-based speaking tasks as part of homework or independent study, allowing students to extend their practice beyond the classroom. To ensure successful implementation, educators should be supported in developing their own digital literacy skills and in learning how to effectively guide and monitor students' AI-assisted activities within the curriculum.

These recommendations, grounded in the outcomes of the current research, aim to support the effective and responsible use of AI technologies in foreign language instruction, contributing to both the linguistic and technological development of students in EFL contexts.

REFERENCES

Al Mahrooqi, Z. A. S. (2024). An Examination of Chatgpt's Efficacy on the 10th Graders' English-Speaking Skills and Motivation Toward English Learning in Oman. (Publication No. 31972864) [Master's thesis, Sultan Qaboos University]. ProQuest Dissertations & Theses Global.

Almineeai, A. M., Airashed, M. A., Hezam, T. A., Alharthi, A. F., & Alharthi, S. F. (2025). The Impact of ChatGPT on EFL Medical Track Students' Affective Filter. *Journal of Language Teaching and Research*, *16*(1), 158-167. https://doi.org/10.17507/jltr.1601.17.

Apriani, E., Cardoso, L., Obaid, A. J., Muthmainnah, M., Wijayanti, E., Esmianti, F., & Supardan, D. (2024). Impact of AI-Powered ChatBots on EFL Students' Writing Skills, Self-Efficacy, and Self-Regulation: A Mixed-Methods Study. *Global Educational Research Review*, *1*(2), 57-72. https://doi.org/10.71380/gerr-08-2024-8

Ballıdağ, M., & Aydın, S. (2025). A comparison of the effects of AI-based chatbots and peer interactions on speaking anxiety among EFL learners. *Future in Educational Research*, *3*(2), 224-238. https://doi.org/10.1002/fer3.70005.

Baskara, F. R. (2025). ChatGPT and Google Gemini in EFL Education: A Qualitative Exploration of Pedagogical Efficacy among Indonesian Sophomores. *Journal of Languages and Language Teaching*, *13*(1), 436-447. https://doi.org/10.33394/jollt.v13i1.9926

Busso, A., & Sanchez, B. (2024). Advancing communicative competence in the digital age: A case for AI tools in Japanese EFL education. *Technology in Language Teaching & Learning*, *6*(3), 1211-1211. https://doi.org/10.29140/tltl.v6n3.1211.

Campbell, D. T. (1963). Experimental and quasi-experimental designs for research on teaching. *Handbook of research on teaching*, *5*, 171-246.

Celik, B., Yildiz, Y., & Kara, S. (2025). Using ChatGPT as a virtual speaking tutor to boost EFL learners' speaking self-efficacy. *Australian Journal of Applied Linguistics*, *8*(1), 102418. https://doi.org/10.29140/ajal.v8n1.102418

Dang, T. P. (2024). Empirical research trends in the use of AI chatbots in EFL teaching and learning in Vietnam: A systematic review. In *Proceedings of the AsiaCALL International Conference* (Vol. 6, pp. 325-343). https://doi.org/10.54855/paic.24623.

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of research in personality*, *19*(2), 109-134.

Erickson, F. (2011). A history of qualitative inquiry in social and educational research. *The Sage handbook of qualitative research*, *4*, 43-59.

Fischer, H. E., Boone, W. J., & Neumann, K. (2023). Quantitative research designs and approaches. In *Handbook of research on science education* (pp. 28-59). Routledge. https://doi.org/10.4324/9780203097267.CH2.

Gardner, R. C. (2014). Attitudes and motivation in second language learning. In *Bilingualism, multiculturalism, and second language learning* (pp. 43-63). Psychology Press.

Hamutoğlu, N. B., Güngören, Ö. C., Uyanık, G. K., & Erdoğan, D. G. (2017). Dijital okuryazarlık ölçeği: Türkçe'ye uyarlama çalışması. *Ege Eğitim Dergisi*, *18*(1), 408-429.

Han, D. E. (2020). The effects of voice-based AI chatbots on Korean EFL middle school students' speaking competence and affective domains. *Asia-pacific Journal of Convergent Research Interchange*, *6*(7), 71-80. https://doi.org/10.47116/apjcri.2020.07.07.

Hapsari, I. P., & Wu, T. T. (2022, August). AI chatbots learning model in English speaking skill: Alleviating speaking anxiety, boosting enjoyment, and fostering critical thinking. In *International conference on innovative technologies and learning* (pp. 444-453). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-15273-3_49

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern language journal*, *70*(2), 125-132.

Kayaoğlu, M. N., & Sağlamel, H. (2013). Students' perceptions of language anxiety in speaking classes. *Journal of History Culture and Art Research*, *2*(2), 142-160.

Krajka, J., & Olszak, I. (2024). AI, will you help?. *How learners use Artificial Intelligence when writing. XLinguae*, *17*(1), 34-48 https://doi.org/10.18355/xl.2024.17.01.05

Krashen, S. D. (1985). The input hypothesis: Issues and implications. Longman Group UK ltd.

Lee, K. A., Lim, S. B., & Nagarajan, S. N. (2023, February). A Study of the Effectiveness of English Speaking of Teachable Agent using AI Chatbot. In *ICAART (1)* (pp. 308-314). https://doi.org/10.5220/0011730300003393

Liu, G. L., Darvin, R., & Ma, C. (2024). Exploring AI-mediated informal digital learning of English (AI-IDLE): A mixed-method investigation of Chinese EFL learners' AI adoption and experiences. *Computer Assisted Language Learning*, 1-29. https://doi.org/10.54097/1g7q6h12

Long, M. (1996). The role of the linguistic environment in second language acquisition. *Handbook of second language acquisition*.

Mingyan, M., Noordin, N., & Razali, A. B. (2024). Effects of an AI Chatbot Mobile Application on Foreign Language Anxiety among Chinese EFL Undergraduates. *International Journal of Academic Research in Progressive Education and Development,* 13(4), 3828–3839. https://doi.org/10.6007/ijarped/v13-i4/23894

Iswari, M. N. P., Santosa, M. H., & Indrayani, L. (2024). Transforming EFL Instruction: A Systematic Review of Chatbot Integration in English Language Learning. *Jurnal Pendidikan Bahasa Inggris undiksha*, *12*(3), 302-309. https://doi.org/10.23887/jpbi.v12i3.92081

Schmidt, R., Boraie, D., & Kassabgy, O. (1996). Foreign language motivation: Internal structure and external connections. Language Learning Motivation: Pathways to the New Century. 14. 9-70.

Shahriar, A. (2025). Exploring efl students' perspectives on the use of ai chatbots in language education. *Journal of Language and Linguistics in Society,* 5, 1-11 DOI: 10.55529/jlls.52.1.11

Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems*, *38*(3), e12667. https://doi.org/10.1111/exsy.12667

Streiner, D. (2003). Starting at the Beginning: An Introduction to Coefficient Alpha and Internal Consistency. *Journal of Personality Assessment*, 80, 103 - 99. https://doi.org/10.1207/S15327752JPA8001_18

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (Vol. 86). Harvard university press. https://doi.org/10.2307/j.ctvjf9vz4

Wei, W., Zhao, A., & Ma, H. (2025). Understanding How AI Chatbots Influence EFL learners' Oral English Learning Motivation and Outcomes: Evidence from Chinese learners. *IEEE Access*. 13, 56699-56716. https://doi.org/10.1109/ACCESS.2025.3554545.

Yang, H., Kim, H., Lee, J., & Shin, D. (2022). Implementation of an AI chatbot as an English conversation partner in EFL speaking classes. *ReCALL*, 34, 327 - 343. https://doi.org/10.1017/S0958344022000039

Young, J. C., & Shishido, M. (2023). Investigating OpenAI's ChatGPT potentials in generating Chatbot's dialogue for English as a foreign language learning. *International journal of advanced computer science and applications*, *14*(6). https://doi.org/10.14569/ijacsa.2023.0140607

Zhang, M., Wang, Y., & Shang, Q. (2024, September). Using AI chatbots to improve university EFL students' speech scripts' writing quality and critical thinking skills. In *Proceedings of the International CALL Research Conference* (Vol. 2024, pp. 345-354). https://doi.org/10.29140/9780648184485-52

Zhou, Q., Hashim, H., & Sulaiman, N. A. (2025). Supporting English speaking practice in higher education: the impact of AI chatbot-integrated mobile-assisted blended learning framework. *Education and Information Technologies*, 1-32. https://doi.org/10.1007/s10639-025-13401-2

APPENDIX

Appendix 1. Digital Literacy Scale

DİJİTAL OKURYAZARLIK ÖLÇEĞİ

Yönerge:

Aşağıdaki ifadeleri dikkatlice okuyunuz ve sizin için ne derece geçerli olduğuna **1'den 5'e kadar** bir puan vererek cevaplayınız.

1 - Kesinlikle Katılmıyorum
2 - Katılmıyorum
3 - Kararsızım
4 - Katılıyorum
5 - Kesinlikle Katılıyorum

	Madde	1	2	3	4	5
1.	Öğrenme sürecinde bilgi ve iletişim teknolojilerini kullanmak hoşuma gider.	☐	☐	☐	☐	☐
2.	Bilgi ve iletişim teknolojilerini kullanarak daha iyi öğrenirim.	☐	☐	☐	☐	☐
3.	Bilgi ve iletişim teknolojilerini kullanarak öğrenmek daha ilgi çekicidir.	☐	☐	☐	☐	☐
4.	Bilgi ve iletişim teknolojilerini kullanarak öğrenmek beni daha motive eder.	☐	☐	☐	☐	☐
5.	Öğrenme etkinliklerim için arkadaşlarımdan sıklıkla internet aracılığıyla yardım alırım.	☐	☐	☐	☐	☐
6.	Bilgi ve iletişim teknolojilerini kullanarak öğrenmek, öz-yönetimli ve bağımsız olmamı sağlar.	☐	☐	☐	☐	☐
7.	Karşılaştığım teknik problemleri nasıl çözeceğimi bilirim.	☐	☐	☐	☐	☐
8.	Yeni teknolojilerin kullanımını kolaylıkla öğrenebilirim.	☐	☐	☐	☐	☐
9.	Önemli olduğunu düşündüğüm yeni teknolojilere ayak uydurabilirim.	☐	☐	☐	☐	☐
10.	Birçok farklı teknoloji hakkında bilgim var.	☐	☐	☐	☐	☐
11.	Öğrenmede ve yeni şeyler oluşturmada (Sunumlar, dijital hikayeler, wikiler, bloglar vb.) bilgi ve iletişim teknolojilerini kullanmak için gerekli olan teknik becerilere sahibim.	☐	☐	☐	☐	☐
12.	İnternetten bilgi elde etmeye yönelik araştırma ve değerlendirme becerilerime güvenirim.	☐	☐	☐	☐	☐
13.	Öğrenme sürecinde mobil teknolojilerin (Cep telefonları, tabletler, akıllı telefonlar vb.) kullanım potansiyeli yüksektir.	☐	☐	☐	☐	☐

Madde	1	2	3	4	5
14. Öğretmenlerim ders anlatırken bilgi ve iletişim teknolojilerini daha çok kullanmalıdır.	☐	☐	☐	☐	☐
15. Bilgi ve iletişim teknolojileri, proje çalışmalarında ve diğer öğrenme etkinliklerinde arkadaşlarımla daha iyi işbirliği içinde çalışmamı sağlar.	☐	☐	☐	☐	☐
16. Bilgi ve iletişim teknolojileri becerilerim iyidir.	☐	☐	☐	☐	☐
17. İnternet tabanlı aktivitelerle ilgili konuları (Örneğin; siber güvenlik, eser hırsızlığı, araştırma teknikleri vb.) bilirim.	☐	☐	☐	☐	☐

Appendix 2 Foreign Language Motivation Scale

Yabancı Dil Öğrenme Motivasyonu Anketi

Yönerge:

Aşağıdaki ifadeleri dikkatlice okuyunuz ve sizin için ne derece geçerli olduğuna **1'den 5'e kadar** bir puan vererek cevaplayınız. **1 "Hiç katılmıyorum"**, **5 ise "Tamamen katılıyorum"** anlamına gelmektedir.

İçsel Motivasyon

Madde	1	2	3	4	5
1. İngilizce öğrenmekten çok keyif alıyorum.	☐	☐	☐	☐	☐
2. İngilizce öğrenmek benim için bir hobi.	☐	☐	☐	☐	☐
3. İngilizce öğrenmek benim için keyifli bir meydan okuma.	☐	☐	☐	☐	☐
4. İngilizce öğrenmekten hoşlanmıyorum, ancak İngilizce öğrenmenin benim için önemli olduğunu biliyorum.	☐	☐	☐	☐	☐
5. Keşke İngilizceyi sınıfa gitmeden daha kolay bir şekilde öğrenebilseydim.	☐	☐	☐	☐	☐

Dışsal Motivasyon

Madde	1	2	3	4	5
6. İngilizce benim için önemli çünkü bakış açımı genişletecek.	☐	☐	☐	☐	☐
7. Bu dersi almamın temel nedeni, ailem/eşim/yöneticilerimin İngilizcemi geliştirmemi istemesi.	☐	☐	☐	☐	☐
8. Bu derste başarılı olmak istiyorum çünkü aileme/arkadaşlarıma/yöneticilerime/başkalarına yeteneğimi göstermek benim için önemli.	☐	☐	☐	☐	☐

Madde	1	2	3	4	5
9. Türkiye'de herkesin İngilizce konuşabilmesi gerekir.	☐	☐	☐	☐	☐
10. İngilizce konuşabilmek sosyal statümü artıracak.	☐	☐	☐	☐	☐
11. İngilizce öğreniyorum çünkü İngilizce konuşulan bir ülkede bir süre yaşamak istiyorum.	☐	☐	☐	☐	☐
12. İngilizce öğrenmek istiyorum çünkü seyahat ederken çok faydalı oluyor.	☐	☐	☐	☐	☐
13. İngilizce öğrenmek istiyorum çünkü yurt dışına göç etmeyi düşünüyorum.	☐	☐	☐	☐	☐
14. İngilizce öğrenme nedenlerimden biri de yeni insanlarla tanışmak ve İngilizce sınıfımdaki arkadaşlar edinmek.	☐	☐	☐	☐	☐
15. Daha eğitimli bir insan olmak için İngilizce öğreniyorum.	☐	☐	☐	☐	☐
16. Ders kitaplarını okuyabilmek için İngilizce bilmem gerekiyor.	☐	☐	☐	☐	☐
17. İngilizce öğrenmemin temel sebebi sınavları geçmek.	☐	☐	☐	☐	☐
18. İngilizceyi daha iyi öğrenirsem, daha iyi bir iş bulabilirim.	☐	☐	☐	☐	☐
19. İngilizce becerilerimi artırmam bana maddi fayda sağlayacak.	☐	☐	☐	☐	☐
20. İngilizce konuşabilirsem harika bir hayatım olacak.	☐	☐	☐	☐	☐

Kişisel Hedefler

Madde	1	2	3	4	5
21. Bu derste, geçmişe göre daha fazla İngilizce öğrenmek istiyorum.	☐	☐	☐	☐	☐
22. Sınıfımdaki diğer öğrencilerden daha başarılı olmak benim için önemli.	☐	☐	☐	☐	☐
23. Bu dersteki öğretmenimle olan ilişkim benim için önemli.	☐	☐	☐	☐	☐

Madde	1	2	3	4	5
24. Bu derste en önemli şeylerden biri diğer öğrencilerle iyi anlaşmak.	☐	☐	☐	☐	☐
25. Bu ders benim için önemli çünkü İngilizceyi iyi öğrenirsem, çocuklarımın da İngilizce öğrenmesine yardımcı olabilirim.	☐	☐	☐	☐	☐

Beklenti/Kontrol Bileşenleri

Madde	1	2	3	4	5
26. Bu İngilizce dersi kesinlikle İngilizcemi geliştirmeme yardımcı olacak.	☐	☐	☐	☐	☐
27. Bu derste başarılı olursam, bunun nedeni çok çalışmam olacak.	☐	☐	☐	☐	☐
28. Bu derste başarılı olmayı bekliyorum çünkü İngilizce öğrenmede iyiyim.	☐	☐	☐	☐	☐
29. Bu derste başarılı olamazsam, bunun nedeni yeterince çaba göstermememem olacak.	☐	☐	☐	☐	☐
30. Bu derste başarılı olamazsam, bunun nedeni İngilizce öğrenme yeteneğimin düşük olması olacak.	☐	☐	☐	☐	☐
31. Bu derste çok şey öğrenirsem, bunun sebebi öğretmen olacak.	☐	☐	☐	☐	☐
32. Bu derste başarılı olursam, bunun nedeni dersin kolay olması olacak.	☐	☐	☐	☐	☐
33. Eğer bu derste iyi öğrenemezsem, bunun temel sebebi öğretmen olacak.	☐	☐	☐	☐	☐
34. Eğer bu derste başarılı olamazsam, bunun sebebi dersin çok zor olması olacak.	☐	☐	☐	☐	☐

Tutumlar

Madde	1	2	3	4	5
35. Amerikalılar çok arkadaş canlısı insanlardır.	☐	☐	☐	☐	☐

Madde	1	2	3	4	5
36. İngilizler geleneklerine bağlı ve muhafazakar insanlardır.	☐	☐	☐	☐	☐
37. En sevdiğim aktörler ve müzisyenlerin çoğu ya İngiliz ya da Amerikalı.	☐	☐	☐	☐	☐
38. İngiliz kültürü dünyaya çok şey katmıştır.	☐	☐	☐	☐	☐

Kaygı

Madde	1	2	3	4	5
39. İngilizce sınıfında konuşmam gerektiğinde kendimi rahatsız hissediyorum.	☐	☐	☐	☐	☐
40. İngilizce sınıfında gönüllü olarak cevap vermek beni utandırıyor.	☐	☐	☐	☐	☐
41. İngilizce sınıfında sık sık konuşmayı sevmiyorum çünkü öğretmenimin beni kötü bir öğrenci olarak düşüneceğinden korkuyorum.	☐	☐	☐	☐	☐
42. İngilizce konuştuğumda diğer öğrencilerin bana güleceğinden korkuyorum.	☐	☐	☐	☐	☐
43. İngilizceyi iyi öğrenebileceğimi düşünüyorum, ancak sınavlarda ve testlerde iyi performans gösteremiyorum.	☐	☐	☐	☐	☐
44. İngilizce sınıfında dikkatim dağılmakta sık sık zorlanıyorum.	☐	☐	☐	☐	☐

Motivasyonel Güç

Madde	1	2	3	4	5
45. Bu dersin ücreti artırılsa bile, İngilizce çalışmak benim için önemli olduğu için kaydolurdum.	☐	☐	☐	☐	☐
46. Bu derse devamım iyi olacak.	☐	☐	☐	☐	☐
47. İngilizceyi mümkün olduğunca uzun süre çalışmaya devam etmeyi planlıyorum.	☐	☐	☐	☐	☐

Madde	1	2	3	4	5
48. Bu dersi bitirdikten sonra muhtemelen başka bir İngilizce dersi alacağım.	☐	☐	☐	☐	☐
49. İngilizceyi nasıl daha iyi öğrenebileceğimi sık sık düşünüyorum.	☐	☐	☐	☐	☐
50. İngilizce öğrenmeye çalışırken gerçekten elimden gelen en iyi çabayı gösterdiğimi dürüstçe söyleyebilirim.	☐	☐	☐	☐	☐

Öğretim Yöntemi Tercihleri

Madde	1	2	3	4	5
51. İngilizce dersinde sadece İngilizce konuşulmasını isterim.	☐	☐	☐	☐	☐
52. İngilizce dersinde öğretmen bazen Türkçe açıklamalar yaparak öğrenmemize yardımcı olmalı.	☐	☐	☐	☐	☐
53. İngilizce dersinde öğretmenin disiplini sağlaması önemlidir.	☐	☐	☐	☐	☐
54. İngilizce dersinde öğretmen konuşmalı, öğrenciler ise sadece çağrıldıklarında cevap vermelidir.	☐	☐	☐	☐	☐
55. Öğrenciler, neden İngilizce öğrendiklerini öğretmene söylemelidir, böylece dersler onların hedeflerine uygun hale getirilebilir.	☐	☐	☐	☐	☐
56. Öğrenciler, anlamadıkları her noktada soru sormalıdır.	☐	☐	☐	☐	☐
57. Öğrencilerin çiftler veya küçük gruplar halinde çalıştığı İngilizce öğrenme etkinliklerini seviyorum.	☐	☐	☐	☐	☐
58. İngilizce dersinde tek başıma çalışmayı, diğer öğrencilerle çalışmaya tercih ederim.	☐	☐	☐	☐	☐
59. İngilizce dersinde grup çalışması ve çiftler halinde çalışma zaman kaybıdır.	☐	☐	☐	☐	☐
60. Öğretmen, bu sınıftaki herkesin eşit derecede iyi İngilizce öğrenmesini sağlamalıdır.	☐	☐	☐	☐	☐
61. İngilizce dersinde en önemli şey dilbilgisidir.	☐	☐	☐	☐	☐

Madde	1	2	3	4	5
62. Telaffuz İngilizce dersinde önemli bir odak noktası olmamalıdır.	☐	☐	☐	☐	☐
63. İngilizce dersinde okuma ve yazma vurgulanmalıdır.	☐	☐	☐	☐	☐
64. İngilizce dersinde dinleme ve konuşma vurgulanmalıdır.	☐	☐	☐	☐	☐
65. Bu derste yapılan etkinlikler, öğrencilerin İngilizce iletişim becerilerini geliştirmelerine yardımcı olacak şekilde tasarlanmalıdır.	☐	☐	☐	☐	☐
66. İletişim etkinlikleri zaman kaybıdır çünkü benim sadece sınavları geçmek için gerekenleri öğrenmem yeterlidir.	☐	☐	☐	☐	☐
67. Bu gibi derslerde, daha fazla öğrenebilmem için beni zorlayan etkinlikler ve materyalleri tercih ederim.	☐	☐	☐	☐	☐
68. İngilizce dersinde ilgimi çeken materyaller kullanılmasını isterim, zor bile olsalar.	☐	☐	☐	☐	☐
69. Aktif olarak katılmamı sağlayan birçok etkinliğin olduğu bir İngilizce dersini tercih ederim.	☐	☐	☐	☐	☐
70. Sadece oturup dinlemeyi tercih ederim, İngilizce dersinde konuşmaya zorlanmak istemem.	☐	☐	☐	☐	☐
71. Öğretmenin sınıfta anında geri bildirim vermesi önemlidir, böylece öğrenciler doğru ya da yanlış yaptıklarını hemen anlayabilirler.	☐	☐	☐	☐	☐
72. Öğretmen, sınıfta hata yapan öğrencileri eleştirmemelidir.	☐	☐	☐	☐	☐

Appendix 3 Foreign Language Speaking Anxiety Scale

Yabancı Dil Konuşma Kaygısı Ölçeği

Yönerge:

Aşağıdaki ifadeleri dikkatlice okuyunuz ve sizin için ne derece geçerli olduğuna **1'den 5'e kadar** bir puan vererek cevaplayınız.

1 - Hiç katılmıyorum
2 - Katılmıyorum
3 - Kararsızım
4 - Katılıyorum
5 - Tamamen katılıyorum

Bölüm 1: Konuşma Kaygısı

Madde	1	2	3	4	5
1. Konuşma dersimde kendimden hiçbir zaman tam olarak emin olamıyorum.	☐	☐	☐	☐	☐
2. Konuşma dersinde hata yapma konusunda endişelenmem.	☐	☐	☐	☐	☐
3. Konuşma dersinde söz hakkı alacağımı bildiğimde titremeye başlıyorum.	☐	☐	☐	☐	☐
4. Öğretmenin söylediği şeyleri anlamadığımda korkuyorum.	☐	☐	☐	☐	☐
5. Daha fazla konuşma dersi almak beni rahatsız etmez.	☐	☐	☐	☐	☐
6. Konuşma dersinde, dersle ilgisi olmayan şeyleri düşündüğüm oluyor.	☐	☐	☐	☐	☐
7. Diğer öğrencilerin benden daha iyi olduğunu düşünüyorum.	☐	☐	☐	☐	☐

Bölüm 2: Performans Kaygısı

Madde	1	2	3	4	5
8. Hazırlık yapmadan konuşmam gerektiğinde panik yapıyorum.	☐	☐	☐	☐	☐
9. Konuşma dersinde başarısız olmanın sonuçları beni endişelendiriyor.	☐	☐	☐	☐	☐
10. Bazı insanların neden konuşma dersinde çok üzüldüğünü anlamıyorum.	☐	☐	☐	☐	☐
11. Konuşma dersinde o kadar heyecanlanıyorum ki bildiğim şeyleri unutuyorum.	☐	☐	☐	☐	☐
12. Konuşma dersinde gönüllü cevap vermek beni utandırıyor.	☐	☐	☐	☐	☐

Madde	1	2	3	4	5
13. Ana dili İngilizce olan kişilerle konuşurken gergin olmam.	☐	☐	☐	☐	☐

Bölüm 3: Konuşma Ortamındaki Stres

Madde	1	2	3	4	5
14. Öğretmenin yaptığı düzeltmeleri anlamadığımda üzülüyorum.	☐	☐	☐	☐	☐
15. Konuşma dersine iyi hazırlandığım halde yine de endişeleniyorum.	☐	☐	☐	☐	☐
16. Çoğu zaman konuşma dersine gitmek istemiyorum.	☐	☐	☐	☐	☐
17. Konuşma dersinde kendimi özgüvenli hissediyorum.	☐	☐	☐	☐	☐
18. Öğretmenimin her hatamı düzeltmeye hazır olduğunu düşünüyorum.	☐	☐	☐	☐	☐
19. Konuşma dersinde söz hakkı alacağım zaman kalbimin hızlı attığını hissediyorum.	☐	☐	☐	☐	☐

Bölüm 4: Sosyal Karşılaştırma ve Özgüven

Madde	1	2	3	4	5
20. Konuşma dersine çok iyi hazırlanmak zorunda hissetmiyorum.	☐	☐	☐	☐	☐
21. Diğer öğrencilerin benden daha iyi konuştuğunu her zaman hissediyorum.	☐	☐	☐	☐	☐
22. Diğer öğrencilerin önünde yabancı dilde konuşurken kendimi çok gergin hissediyorum.	☐	☐	☐	☐	☐
23. Konuşma dersi çok hızlı ilerlediği için geri kalmaktan endişe duyuyorum.	☐	☐	☐	☐	☐
24. Konuşma dersim diğer derslerime göre daha fazla stres ve gerginlik hissetmeme neden oluyor.	☐	☐	☐	☐	☐
25. Konuşma dersinde konuşurken kafam karışıyor ve gergin oluyorum.	☐	☐	☐	☐	☐

Bölüm 5: Konuşma Kaygısının Günlük Yaşama Etkisi

Madde	1	2	3	4	5
26. Konuşma dersine giderken kendimi çok rahat ve özgüvenli hissediyorum.	☐	☐	☐	☐	☐
27. Konuşma öğretmenimin söylediği her kelimeyi anlamadığımda gergin oluyorum.	☐	☐	☐	☐	☐

Madde	1	2	3	4	5
28. Yabancı bir dili konuşabilmek için öğrenilmesi gereken kuralların fazlalığı beni bunaltıyor.	☐	☐	☐	☐	☐
29. Diğer öğrencilerin benim konuşmamla dalga geçmesinden korkuyorum.	☐	☐	☐	☐	☐
30. Ana dili İngilizce olan kişilerle konuşurken kendimi rahat hissederim.	☐	☐	☐	☐	☐
31. Öğretmenim önceden hazırlamadığım sorular sorduğunda gergin oluyorum.	☐	☐	☐	☐	☐

Appendix 4 Semi Structured Interview Questions

Yarı Yapılandırılmış Görüşme Soruları

1. Genel Deneyim

1.1. Bu süreçte yapay zeka destekli konuşma etkinliklerine katılmak sizin için nasıl bir deneyimdi?

1.2. Bu etkinlikler sizin konuşma becerilerinizi geliştirme açısından nasıl bir katkı sağladı?

2. Kaygı ve Konfor Düzeyi

2.1. Yapay zeka ile konuşurken kendinizi rahat hissettiniz mi? Neden?

2.2. Gerçek bir öğretmenle konuşmakla karşılaştırıldığında bu deneyim sizi ne yönde etkiledi?

3. Motivasyon ve Katılım

3.1. Yapay zeka ile yapılan etkinlikler, İngilizce konuşma konusunda sizi motive etti mi? Nasıl?

3.2. Ders dışında da bu uygulamaları kullanmak istediniz mi? Neden?

4. Teknolojik Uygulama ve Kullanım Kolaylığı

4.1. Uygulamayı veya sistemi kullanırken teknik bir sorun yaşadınız mı?

4.2. Bu tür bir uygulamanın sınıf içi eğitime entegre edilmesini nasıl değerlendiriyorsunuz?

5. Tercih ve Gelecek Uygulamalar

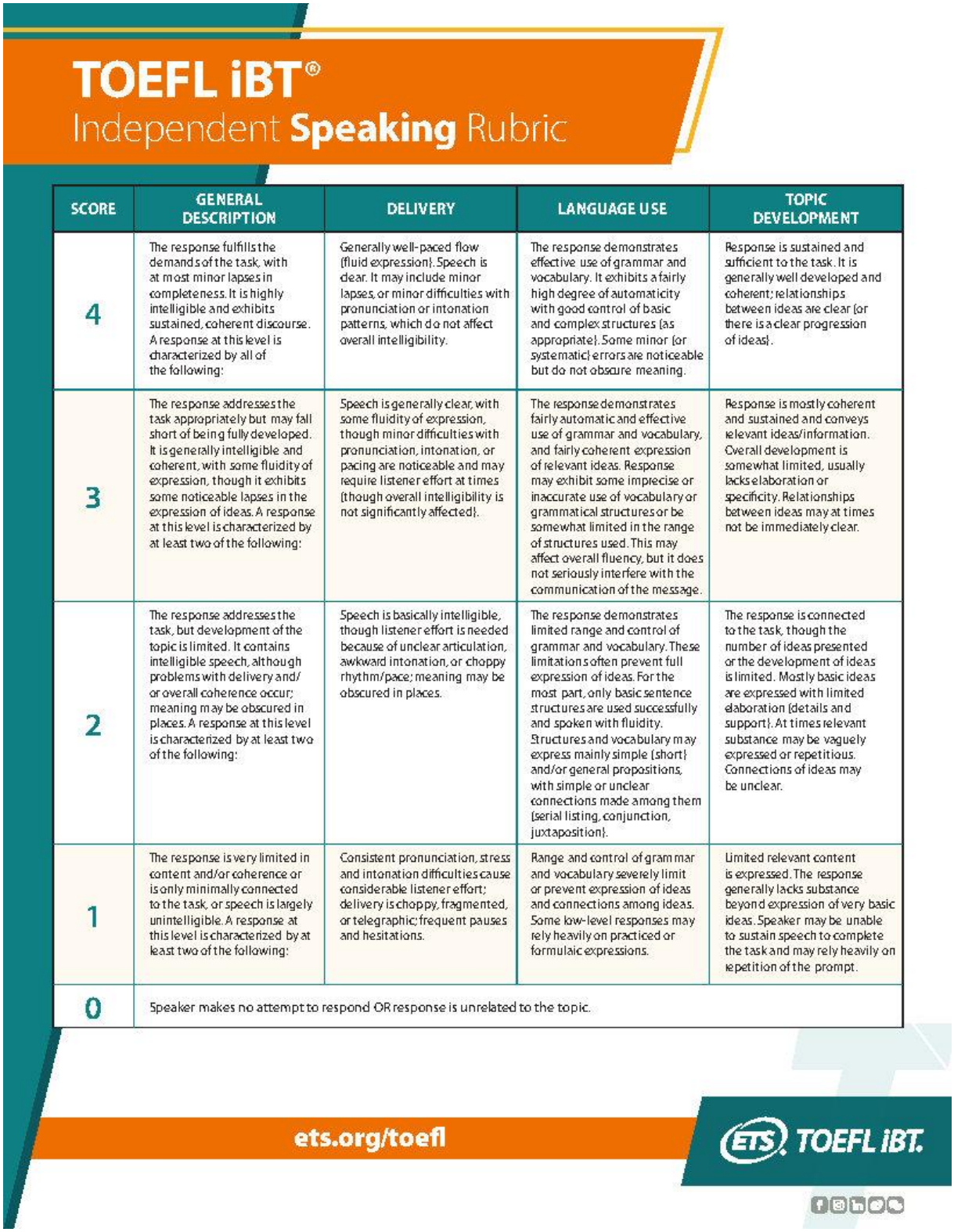
5.1. Gelecekte İngilizce konuşma pratiği için yapay zeka destekli sistemleri kullanmak ister misiniz?

5.2. Eğer sizin kararınız olsaydı, bu sistemi nasıl geliştirirdiniz?

6. Ek Görüşler

6.1. Bu süreçle ilgili eklemek istediğiniz başka bir görüşünüz veya paylaşmak istediğiniz bir deneyiminiz var mı?

Appendix 5 TOEFL IBT Independent Speaking Rubric



TOEFL iBT®
Independent **Speaking** Rubric

SCORE	GENERAL DESCRIPTION	DELIVERY	LANGUAGE USE	TOPIC DEVELOPMENT
4	The response fulfills the demands of the task, with at most minor lapses in completeness. It is highly intelligible and exhibits sustained, coherent discourse. A response at this level is characterized by all of the following:	Generally well-paced flow (fluid expression). Speech is clear. It may include minor lapses, or minor difficulties with pronunciation or intonation patterns, which do not affect overall intelligibility.	The response demonstrates effective use of grammar and vocabulary. It exhibits a fairly high degree of automaticity with good control of basic and complex structures (as appropriate). Some minor (or systematic) errors are noticeable but do not obscure meaning.	Response is sustained and sufficient to the task. It is generally well developed and coherent; relationships between ideas are clear (or there is a clear progression of ideas).
3	The response addresses the task appropriately but may fall short of being fully developed. It is generally intelligible and coherent, with some fluidity of expression, though it exhibits some noticeable lapses in the expression of ideas. A response at this level is characterized by at least two of the following:	Speech is generally clear, with some fluidity of expression, though minor difficulties with pronunciation, intonation, or pacing are noticeable and may require listener effort at times (though overall intelligibility is not significantly affected).	The response demonstrates fairly automatic and effective use of grammar and vocabulary, and fairly coherent expression of relevant ideas. Response may exhibit some imprecise or inaccurate use of vocabulary or grammatical structures or be somewhat limited in the range of structures used. This may affect overall fluency, but it does not seriously interfere with the communication of the message.	Response is mostly coherent and sustained and conveys relevant ideas/information. Overall development is somewhat limited, usually lacks elaboration or specificity. Relationships between ideas may at times not be immediately clear.
2	The response addresses the task, but development of the topic is limited. It contains intelligible speech, although problems with delivery and/or overall coherence occur; meaning may be obscured in places. A response at this level is characterized by at least two of the following:	Speech is basically intelligible, though listener effort is needed because of unclear articulation, awkward intonation, or choppy rhythm/pace; meaning may be obscured in places.	The response demonstrates limited range and control of grammar and vocabulary. These limitations often prevent full expression of ideas. For the most part, only basic sentence structures are used successfully and spoken with fluidity. Structures and vocabulary may express mainly simple (short) and/or general propositions, with simple or unclear connections made among them (serial listing, conjunction, juxtaposition).	The response is connected to the task, though the number of ideas presented or the development of ideas is limited. Mostly basic ideas are expressed with limited elaboration (details and support). At times relevant substance may be vaguely expressed or repetitious. Connections of ideas may be unclear.
1	The response is very limited in content and/or coherence or is only minimally connected to the task, or speech is largely unintelligible. A response at this level is characterized by at least two of the following:	Consistent pronunciation, stress and intonation difficulties cause considerable listener effort; delivery is choppy, fragmented, or telegraphic; frequent pauses and hesitations.	Range and control of grammar and vocabulary severely limit or prevent expression of ideas and connections among ideas. Some low-level responses may rely heavily on practiced or formulaic expressions.	Limited relevant content is expressed. The response generally lacks substance beyond expression of very basic ideas. Speaker may be unable to sustain speech to complete the task and may rely heavily on repetition of the prompt.
0	Speaker makes no attempt to respond OR response is unrelated to the topic.			

ets.org/toefl

ETS TOEFL iBT.

Appendix 6 Item Analysis of Findings of Motivation Scale Pre-Test Results

Findings of Motivation Scale Pre-Test Results

	Totally Disagree		Disagree		Neutral		Agree		Totally Agree		N	M	SD
	n	%	n	%	n	%	n	%	n	%			
Item 1	3	5	9	15	21	35	11	18.3	16	26.7	60	3.47	1.19
Item 2	4	6.7	9	15	15	25	19	31.7	13	21.7	60	3.47	1.19
Item 3	0	0	10	16.7	15	25	24	40	11	18.3	60	3.6	0.98
Item 4	3	5	17	28.3	9	15	13	21.7	18	30	60	3.43	1.32
Item 5	5	8.3	9	15	15	25	17	28.3	14	23.3	60	3.43	1.24
Item 6	2	3.3	10	16.7	13	21.7	19	31.7	16	26.7	60	3.62	1.15
Item 7	5	8.3	11	18.3	6	10	15	25	23	38.3	60	3.67	1.37
Item 8	2	3.3	14	23.3	16	26.7	12	20	16	26.7	60	3.43	1.21
Item 9	1	1.7	12	20	14	23.3	15	25	18	30	60	3.62	1.17
Item 10	6	10	5	8.3	18	30	19	31.7	12	20	60	3.43	1.23
Item 11	2	3.3	10	16.7	8	13.3	25	41.7	15	25	60	3.68	1.13
Item 12	3	5	5	8.3	15	25	14	23.3	23	38.3	60	3.82	1.19
Item 13	4	6.7	8	13.3	13	21.7	21	35	14	23.3	60	3.55	1.19
Item 14	2	3.3	14	23.3	16	26.7	13	21.7	15	25	60	3.42	1.22
Item 15	3	5	9	15	14	23.3	21	35	13	21.7	60	3.53	1.14
Item 16	3	5	9	15	13	21.7	13	21.7	22	36.7	60	3.7	1.25
Item 17	4	6.7	10	16.7	11	18.3	20	33.3	15	25	60	3.53	1.23

Item 18	3	5	11	18.3	9	15	22	36.7	15	25	60	3.58	1.2
Item 19	1	1.7	11	18.3	18	30	15	25	15	25	60	3.53	1.11
Item 20	3	5	7	11.7	13	21.7	15	25	22	36.7	60	3.77	1.21
Item 21	5	8.3	8	13.3	18	30	15	25	14	23.3	60	3.42	1.23
Item 22	4	6.7	8	13.3	14	23.3	17	28.3	17	28.3	60	3.58	1.23
Item 23	2	3.3	12	20	19	31.7	20	33.3	7	11.7	60	3.3	1.03
Item 24	5	8.3	8	13.3	15	25	17	28.3	15	25	60	3.48	1.24
Item 25	5	8.3	9	15	17	28.3	17	28.3	12	20	60	3.37	1.21
Item 26	3	5	11	18.3	15	25	18	30	13	21.7	60	3.45	1.17
Item 27	3	5	11	18.3	10	16.7	17	28.3	19	31.7	60	3.63	1.25
Item 28	5	8.3	11	18.3	9	15	9	15	26	43.3	60	3.67	1.41
Item 29	1	1.7	9	15	17	28.3	16	26.7	17	28.3	60	3.65	1.1
Item 30	7	11.7	13	21.7	11	18.3	15	25	14	23.3	60	3.27	1.35
Item 31	4	6.7	13	21.7	14	23.3	11	18.3	18	30	60	3.43	1.31
Item 32	6	10	11	18.3	13	21.7	18	30	12	20	60	3.32	1.27
Item 33	5	8.3	6	10	14	23.3	16	26.7	19	31.7	60	3.63	1.26
Item 34	1	1.7	7	11.7	19	31.7	19	31.7	14	23.3	60	3.63	1.02
Item 35	6	10	8	13.3	8	13.3	17	28.3	21	35	60	3.65	1.35
Item 36	3	5	8	13.3	15	25	12	20	22	36.7	60	3.7	1.24
Item 37	7	11.7	7	11.7	13	21.7	18	30	15	25	60	3.45	1.31
Item 38	0	0	7	11.7	16	26.7	21	35	16	26.7	60	3.77	0.98

Item 39	4	6.7	11	18.3	8	13.3	16	26.7	21	35	60	3.65	1.31
Item 40	4	6.7	9	15	20	33.3	15	25	12	20	60	3.37	1.16
Item 41	1	1.7	8	13.3	16	26.7	19	31.7	16	26.7	60	3.68	1.07
Item 42	7	11.7	9	15	14	23.3	17	28.3	13	21.7	60	3.33	1.33
Item 43	5	8.3	4	6.7	19	31.7	20	33.3	12	20	60	3.5	1.14
Item 44	3	5	12	20	16	26.7	13	21.7	16	26.7	60	3.45	1.23
Item 45	6	10	9	15	17	28.3	13	21.7	15	25	60	3.37	1.29
Item 46	6	10	5	8.3	15	25	16	26.7	18	30	60	3.58	1.28
Item 47	5	8.3	7	11.7	18	30	15	25	15	25	60	3.47	1.23
Item 48	5	8.3	13	21.7	13	21.7	19	31.7	10	16.7	60	3.27	1.22
Item 49	2	3.3	10	16.7	13	21.7	18	30	17	28.3	60	3.63	1.16
Item 50	3	5	7	11.7	12	20	24	40	14	23.3	60	3.65	1.12
Item 51	3	5	10	16.7	18	30	17	28.3	12	20	60	3.42	1.14
Item 52	3	5	9	15	17	28.3	15	25	16	26.7	60	3.53	1.19
Item 53	3	5	10	16.7	14	23.3	17	28.3	16	26.7	60	3.55	1.25
Item 54	3	5	11	18.3	24	40	10	16.7	12	20	60	3.28	1.14
Item 55	3	5	9	15	14	23.3	19	31.7	15	25	60	3.57	1.17
Item 56	7	11.7	10	16.7	13	21.7	16	26.7	14	23.3	60	3.33	1.32
Item 57	3	5	3	5	21	35	13	21.7	20	33.3	60	3.73	1.13
Item 58	3	5	10	16.7	20	33.3	13	21.7	14	23.3	60	3.42	1.17
Item 59	0	0	8	13.3	20	33.3	19	31.7	13	21.7	60	3.62	0.98

Item 60	5	8.3	7	11.7	21	35	14	23.3	13	21.7	60	3.38	1.19
Item 61	2	3.3	4	6.7	19	31.7	17	28.3	18	30	60	3.75	1.07
Item 62	3	5	10	16.7	17	28.3	16	26.7	14	23.3	60	3.47	1.17
Item 63	3	5	8	13.3	24	40	13	21.7	12	20	60	3.38	1.11
Item 64	7	11.7	14	23.3	14	23.3	10	16.7	15	25	60	3.2	1.36
Item 65	3	5	11	18.3	16	26.7	12	20	18	30	60	3.52	1.24
Item 66	4	6.7	5	8.3	17	28.3	15	25	19	31.7	60	3.67	1.2
Item 67	6	10	5	8.3	23	38.3	11	18.3	15	25	60	3.4	1.24
Item 68	5	8.3	11	18.3	19	31.7	9	15	16	26.7	60	3.33	1.28
Item 69	4	6.7	10	16.7	13	21.7	14	23.3	19	31.7	60	3.57	1.28
Item 70	5	8.3	5	8.3	19	31.7	10	16.7	21	35	60	3.62	1.28
Item 71	4	6.7	4	6.7	15	25	17	28.3	20	33.3	60	3.75	1.19
Item 72	2	3.3	6	10	20	33.3	17	28.3	15	25	60	3.62	1.08

Appendix 7 Item Analysis of Findings of Motivation Scale Post-Test Results

Findings of Motivation Scale Post-Test Results

	Totally Disagree		Disagree		Neutral		Agree		Totally Agree		N	M	SD
	n	%	n	%	n	%	n	%	n	%			
Item 1	1	1.7	8	13.3	23	38.3	12	20	16	26.7	60	3.57	1.08
Item 2	4	6.7	6	10	19	31.7	20	33.3	11	18.3	60	3.47	1.11
Item 3	3	5	6	10	15	25	16	26.7	20	33.3	60	3.73	1.18
Item 4	0	0	3	5	14	23.3	25	41.7	18	30	60	3.97	0.86
Item 5	2	3.3	2	3.3	15	25	17	28.3	24	40	60	3.98	1.05
Item 6	1	1.7	4	6.7	11	18.3	24	40	20	33.3	60	3.97	0.97
Item 7	0	0	6	10	15	25	19	31.7	20	33.3	60	3.88	0.99
Item 8	2	3.3	10	16.7	18	30	17	28.3	13	21.7	60	3.48	1.11
Item 9	0	0	3	5	22	36.7	14	23.3	21	35	60	3.88	0.96
Item 10	1	1.7	6	10	23	38.3	13	21.7	17	28.3	60	3.65	1.05
Item 11	2	3.3	7	11.7	22	36.7	14	23.3	15	25	60	3.55	1.11
Item 12	2	3.3	6	10	18	30	20	33.3	14	23.3	60	3.63	1.06
Item 13	2	3.3	3	5	18	30	14	23.3	23	38.3	60	3.88	1.09
Item 14	1	1.7	8	13.3	16	26.7	17	28.3	18	30	60	3.72	1.09
Item 15	1	1.7	10	16.7	19	31.7	15	25	15	25	60	3.55	1.11
Item 16	3	5	7	11.7	18	30	16	26.7	16	26.7	60	3.58	1.15
Item 17	2	3.3	2	3.3	18	30	18	30	20	33.3	60	3.87	1.03

Item 18	2	3.3	7	11.7	16	26.7	14	23.3	21	35	60	3.75	1.16
Item 19	1	1.7	3	5	14	23.3	19	31.7	23	38.3	60	4	0.99
Item 20	1	1.7	6	10	19	31.7	18	30	16	26.7	60	3.7	1.03
Item 21	3	5	4	6.7	24	40	14	23.3	15	25	60	3.57	1.09
Item 22	3	5	9	15	17	28.3	16	26.7	15	25	60	3.52	1.17
Item 23	3	5	7	11.7	16	26.7	10	16.7	24	40	60	3.75	1.24
Item 24	2	3.3	9	15	21	35	11	18.3	17	28.3	60	3.53	1.16
Item 25	2	3.3	9	15	18	30	16	26.7	15	25	60	3.55	1.13
Item 26	3	5	8	13.3	14	23.3	20	33.3	15	25	60	3.6	1.15
Item 27	4	6.7	8	13.3	17	28.3	17	28.3	14	23.3	60	3.48	1.19
Item 28	3	5	9	15	18	30	15	25	15	25	60	3.5	1.17
Item 29	1	1.7	6	10	14	23.3	20	33.3	19	31.7	60	3.83	1.04
Item 30	1	1.7	9	15	11	18.3	22	36.7	17	28.3	60	3.75	1.08
Item 31	0	0	6	10	12	20	24	40	18	30	60	3.9	0.95
Item 32	2	3.3	10	16.7	20	33.3	12	20	16	26.7	60	3.5	1.16
Item 33	4	6.7	6	10	22	36.7	16	26.7	12	20	60	3.43	1.13
Item 34	1	1.7	4	6.7	14	23.3	23	38.3	18	30	60	3.88	0.98
Item 35	2	3.3	10	16.7	18	30	15	25	15	25	60	3.52	1.14
Item 36	2	3.3	6	10	21	35	16	26.7	15	25	60	3.6	1.08
Item 37	2	3.3	8	13.3	12	20	22	36.7	16	26.7	60	3.7	1.11
Item 38	2	3.3	7	11.7	19	31.7	21	35	11	18.3	60	3.53	1.03

Item 39	2	3.3	9	15	24	40	13	21.7	12	20	60	3.4	1.08
Item 40	3	5	7	11.7	17	28.3	18	30	15	25	60	3.58	1.14
Item 41	5	8.3	7	11.7	11	18.3	21	35	16	26.7	60	3.6	1.24
Item 42	2	3.3	7	11.7	13	21.7	19	31.7	19	31.7	60	3.77	1.13
Item 43	3	5	6	10	12	20	22	36.7	17	28.3	60	3.73	1.13
Item 44	1	1.7	4	6.7	17	28.3	21	35	17	28.3	60	3.82	0.98
Item 45	3	5	7	11.7	16	26.7	17	28.3	17	28.3	60	3.63	1.16
Item 46	1	1.7	6	10	13	21.7	21	35	19	31.7	60	3.85	1.04
Item 47	4	6.7	6	10	22	36.7	9	15	19	31.7	60	3.55	1.23
Item 48	0	0	8	13.3	13	21.7	20	33.3	19	31.7	60	3.83	1.03
Item 49	2	3.3	10	16.7	15	25	17	28.3	16	26.7	60	3.58	1.15
Item 50	3	5	7	11.7	17	28.3	13	21.7	20	33.3	60	3.67	1.2
Item 51	3	5	6	10	17	28.3	12	20	22	36.7	60	3.73	1.21
Item 52	0	0	7	11.7	19	31.7	15	25	19	31.7	60	3.77	1.03
Item 53	2	3.3	6	10	19	31.7	17	28.3	16	26.7	60	3.65	1.09
Item 54	2	3.3	7	11.7	12	20	20	33.3	19	31.7	60	3.78	1.12
Item 55	1	1.7	5	8.3	27	45	14	23.3	13	21.7	60	3.55	0.98
Item 56	1	1.7	6	10	15	25	19	31.7	19	31.7	60	3.82	1.05
Item 57	3	5	7	11.7	13	21.7	20	33.3	17	28.3	60	3.68	1.16
Item 58	2	3.3	3	5	17	28.3	26	43.3	12	20	60	3.72	0.96
Item 59	1	1.7	6	10	11	18.3	15	25	27	45	60	4.02	1.1

Item 60	0	0	7	11.7	17	28.3	18	30	18	30	60	3.78	1.01
Item 61	2	3.3	7	11.7	17	28.3	16	26.7	18	30	60	3.68	1.13
Item 62	3	5	6	10	18	30	21	35	12	20	60	3.55	1.08
Item 63	3	5	6	10	23	38.3	13	21.7	15	25	60	3.52	1.13
Item 64	1	1.7	10	16.7	16	26.7	16	26.7	17	28.3	60	3.63	1.12
Item 65	0	0	3	5	26	43.3	11	18.3	20	33.3	60	3.8	0.97
Item 66	3	5	7	11.7	19	31.7	14	23.3	17	28.3	60	3.58	1.17
Item 67	3	5	8	13.3	21	35	12	20	16	26.7	60	3.5	1.17
Item 68	3	5	5	8.3	15	25	17	28.3	20	33.3	60	3.77	1.16
Item 69	2	3.3	5	8.3	22	36.7	16	26.7	15	25	60	3.62	1.06
Item 70	3	5	5	8.3	12	20	29	48.3	11	18.3	60	3.67	1.04
Item 71	4	6.7	8	13.3	18	30	14	23.3	16	26.7	60	3.5	1.21
Item 72	3	5	5	8.3	19	31.7	16	26.7	17	28.3	60	3.65	1.13

Appendix 8 Item Analysis of Findings of Digital Literacy Scale Pre-Test Results

Findings of Digital Literacy Scale Pre-Test Results

	Totally Disagree		Disagree		Neutral		Agree		Totally Agree		N	M	SD
	n	%	n	%	n	%	n	%	n	%			
Item 1	3	5	6	10	18	30	12	20	21	35	60	3.7	1.2
Item 2	1	1.7	11	18.3	12	20	21	35	15	25	60	3.63	1.1
Item 3	2	3.3	11	18.3	10	16.7	13	21.7	24	40	60	3.77	1.25
Item 4	3	5	10	16.7	18	30	11	18.3	18	30	60	3.52	1.23
Item 5	1	1.7	12	20	14	23.3	19	31.7	14	23.3	60	3.55	1.11
Item 6	3	5	6	10	19	31.7	18	30	14	23.3	60	3.57	1.11
Item 7	4	6.7	15	25	11	18.3	17	28.3	13	21.7	60	3.33	1.26
Item 8	2	3.3	10	16.7	13	21.7	14	23.3	21	35	60	3.7	1.21
Item 9	5	8.3	14	23.3	16	26.7	9	15	16	26.7	60	3.28	1.32
Item 10	7	11.7	3	5	14	23.3	16	26.7	20	33.3	60	3.65	1.31
Item 11	3	5	14	23.3	13	21.7	15	25	15	25	60	3.42	1.24
Item 12	4	6.7	6	10	18	30	20	33.3	12	20	60	3.5	1.13
Item 13	2	3.3	6	10	12	20	17	28.3	23	38.3	60	3.88	1.14
Item 14	7	11.7	9	15	13	21.7	16	26.7	15	25	60	3.38	1.33
Item 15	3	5	8	13.3	13	21.7	19	31.7	17	28.3	60	3.65	1.18
Item 16	3	5	11	18.3	15	25	18	30	13	21.7	60	3.45	1.17
Item 17	5	8.3	8	13.3	19	31.7	10	16.7	18	30	60	3.47	1.28

Appendix 9 Item Analysis of Findings of Digital Literacy Scale Post-Test Results

Findings of Digital Literacy Scale Post-Test Results

	Totally Disagree		Disagree		Neutral		Agree		Totally Agree		N	M	SD
	n	%	n	%	n	%	n	%	n	%			
Item 1	1	1.7	7	11.7	7	11.7	32	53.3	13	21.7	60	3.82	0.97
Item 2	2	3.3	9	15	14	23.3	17	28.3	18	30	60	3.67	1.16
Item 3	6	10	8	13.3	10	16.7	22	36.7	14	23.3	60	3.5	1.27
Item 4	1	1.7	11	18.3	16	26.7	13	21.7	19	31.7	60	3.63	1.16
Item 5	3	5	8	13.3	14	23.3	14	23.3	21	35	60	3.7	1.23
Item 6	4	6.7	6	10	12	20	20	33.3	18	30	60	3.7	1.2
Item 7	3	5	7	11.7	15	25	20	33.3	15	25	60	3.62	1.14
Item 8	4	6.7	5	8.3	11	18.3	18	30	22	36.7	60	3.82	1.21
Item 9	4	6.7	10	16.7	10	16.7	21	35	15	25	60	3.55	1.23
Item 10	4	6.7	2	3.3	13	21.7	19	31.7	22	36.7	60	3.88	1.15
Item 11	3	5	7	11.7	18	30	15	25	17	28.3	60	3.6	1.17
Item 12	2	3.3	7	11.7	19	31.7	18	30	14	23.3	60	3.58	1.08
Item 13	7	11.7	4	6.7	19	31.7	18	30	12	20	60	3.4	1.22
Item 14	4	6.7	6	10	15	25	18	30	17	28.3	60	3.63	1.19
Item 15	4	6.7	9	15	10	16.7	16	26.7	21	35	60	3.68	1.28
Item 16	5	8.3	1	1.7	15	25	28	46.7	11	18.3	60	3.65	1.07
Item 17	2	3.3	8	13.3	13	21.7	19	31.7	18	30	60	3.72	1.14

Appendix 10 Item Analysis of Findings of FLSAS Scale Pre-Test Results

Findings of FLSAS Scale Pre-Test Results

	Totally Disagree		Disagree		Neutral		Agree		Totally Agree		N	M	SD
	n	%	n	%	n	%	n	%	n	%			
Item 1	6	10	6	10	16	26.7	18	30	14	23.3	60	3.47	1.24
Item 2	6	10	7	11.7	12	20	20	33.3	15	25	60	3.52	1.27
Item 3	3	5	7	11.7	13	21.7	27	45	10	16.7	60	3.57	1.06
Item 4	3	5	8	13.3	16	26.7	17	28.3	16	26.7	60	3.58	1.17
Item 5	4	6.7	9	15	14	23.3	15	25	18	30	60	3.57	1.25
Item 6	5	8.3	10	16.7	15	25	15	25	15	25	60	3.42	1.27
Item 7	5	8.3	8	13.3	21	35	12	20	14	23.3	60	3.37	1.22
Item 8	1	1.7	9	15	20	33.3	14	23.3	16	26.7	60	3.58	1.09
Item 9	6	10	7	11.7	17	28.3	9	15	21	35	60	3.53	1.35
Item 10	3	5	9	15	14	23.3	18	30	16	26.7	60	3.58	1.18
Item 11	6	10	10	16.7	19	31.7	13	21.7	12	20	60	3.25	1.24
Item 12	4	6.7	11	18.3	15	25	13	21.7	17	28.3	60	3.47	1.27
Item 13	4	6.7	5	8.3	14	23.3	18	30	19	31.7	60	3.72	1.19
Item 14	4	6.7	11	18.3	18	30	10	16.7	17	28.3	60	3.42	1.27
Item 15	5	8.3	13	21.7	17	28.3	7	11.7	18	30	60	3.33	1.34
Item 16	5	8.3	8	13.3	14	23.3	18	30	15	25	60	3.5	1.24
Item 17	4	6.7	6	10	29	48.3	7	11.7	14	23.3	60	3.35	1.15

Item 18	3	5	11	18.3	17	28.3	15	25	14	23.3	60	3.43	1.18
Item 19	1	1.7	8	13.3	10	16.7	18	30	23	38.3	60	3.9	1.12
Item 20	5	8.3	8	13.3	16	26.7	16	26.7	15	25	60	3.47	1.24
Item 21	6	10	9	15	12	20	17	28.3	16	26.7	60	3.47	1.31
Item 22	1	1.7	7	11.7	12	20	21	35	19	31.7	60	3.83	1.06
Item 23	5	8.3	7	11.7	15	25	16	26.7	17	28.3	60	3.55	1.25
Item 24	7	11.7	8	13.3	12	20	19	31.7	14	23.3	60	3.42	1.31
Item 25	3	5	11	18.3	15	25	16	26.7	15	25	60	3.48	1.28
Item 26	1	1.7	8	13.3	17	28.3	20	33.3	14	23.3	60	3.63	1.04
Item 27	3	5	7	11.7	16	26.7	17	28.3	17	28.3	60	3.63	1.16
Item 28	2	3.3	6	10	17	28.3	22	36.7	13	21.7	60	3.63	1.04
Item 29	1	1.7	12	20	14	23.3	16	26.7	17	28.3	60	3.6	1.15
Item 30	6	10	4	6.7	11	18.3	18	30	21	35	60	3.73	1.29
Item 31	8	13.3	8	13.3	8	13.3	20	33.3	16	26.7	60	3.47	1.37

Appendix 11 Item Analysis of Findings of FLSAS Scale Post-Test Results

Findings of FLSAS Scale Post-Test Results

	Totally Disagree		Disagree		Neutral		Agree		Totally Agree		N	M	SD
	n	%	n	%	n	%	n	%	n	%	n	n	%
Item 1	0	0	14	23.3	16	26.7	20	33.3	10	16.7	60	3.43	1.03
Item 2	0	0	13	21.7	25	41.7	12	20	10	16.7	60	3.32	1
Item 3	5	8.3	12	20	21	35	13	21.7	9	15	60	3.15	1.16
Item 4	4	6.7	6	10	16	26.7	23	38.3	11	18.3	60	3.52	1.11
Item 5	5	8.3	12	20	17	28.3	18	30	8	13.3	60	3.2	1.16
Item 6	3	5	8	13.3	17	28.3	23	38.3	9	15	60	3.45	1.06
Item 7	2	3.3	9	15	17	28.3	21	35	11	18.3	60	3.5	1.07
Item 8	8	13.3	11	18.3	13	21.7	16	26.7	12	20	60	3.22	1.33
Item 9	4	6.7	14	23.3	13	21.7	21	35	8	13.3	60	3.25	1.16
Item 10	4	6.7	9	15	17	28.3	16	26.7	14	23.3	60	3.45	1.2
Item 11	4	6.7	11	18.3	18	30	19	31.7	8	13.3	60	3.27	1.12
Item 12	4	6.7	10	16.7	18	30	18	30	10	16.7	60	3.33	1.14
Item 13	3	5	13	21.7	15	25	16	26.7	13	21.7	60	3.38	1.19
Item 14	2	3.3	14	23.3	19	31.7	14	23.3	11	18.3	60	3.3	1.12
Item 15	6	10	6	10	22	36.7	18	30	8	13.3	60	3.27	1.13
Item 16	5	8.3	8	13.3	15	25	19	31.7	13	21.7	60	3.45	1.21
Item 17	3	5	13	21.7	22	36.7	10	16.7	12	20	60	3.25	1.16

Item 18	7	11.7	10	16.7	17	28.3	18	30	8	13.3	60	3.17	1.21
Item 19	4	6.7	14	23.3	10	16.7	19	31.7	13	21.7	60	3.38	1.25
Item 20	8	13.3	10	16.7	14	23.3	16	26.7	12	20	60	3.23	1.32
Item 21	4	6.7	9	15	14	23.3	21	35	12	20	60	3.47	1.17
Item 22	5	8.3	6	10	18	30	19	31.7	12	20	60	3.45	1.17
Item 23	3	5	8	13.3	20	33.3	13	21.7	16	26.7	60	3.52	1.17
Item 24	5	8.3	10	16.7	16	26.7	22	36.7	7	11.7	60	3.27	1.13
Item 25	7	11.7	6	10	22	36.7	15	25	10	16.7	60	3.25	1.2
Item 26	2	3.3	18	30	16	26.7	14	23.3	10	16.7	60	3.2	1.15
Item 27	3	5	15	25	12	20	21	35	9	15	60	3.3	1.15
Item 28	0	0	13	21.7	20	33.3	15	25	12	20	60	3.43	1.05
Item 29	5	8.3	16	26.7	14	23.3	14	23.3	11	18.3	60	3.17	1.25
Item 30	2	3.3	12	20	9	15	23	38.3	14	23.3	60	3.58	1.15
Item 31	5	8.3	11	18.3	20	33.3	15	25	9	15	60	3.2	1.16

Appendix 12 Approval of Ministry of Education

T.C.
MİLLÎ EĞİTİM BAKANLIĞI
ARAŞTIRMA UYGULAMA İZNİ BELGESİ

EĞİTİMİN YÜZYILI

Başvuru No: MEB.TT.2025.023335.01

T.C. Kimlik No: 14*******58

Adı Soyadı: MEHMET TAŞDEMİR

Araştırmanın Adı: Yapay Zeka Sohbet Botlarının EFL Lise Öğrencilerinin Konuşma Becerisi, Kaygı, Motivasyon ve Dijital Okuryazarlık Üzerindeki Etkisinin İncelenmesi

Araştırmanın Niteliği: Yüksek Lisans Tezi

Araştırmanın Örneklem / Çalışma Grubu: Öğrenci

Veri Toplama Aracının Başlığı: DİJİTAL OKURYAZARLIK ÖLÇEĞİ, Yabancı Dil Konuşma Kaygısı Ölçeği, Yabancı Dil Öğrenme Motivasyonu Anketi, Yarı yapılandırılmış görüşme soruları

Araştırma Uygulama İzninin Kabul Tarihi: 18.04.2025

Araştırma Uygulama İzninin Bitiş Tarihi: 18.04.2026

Yukarıda kimliği yazılı araştırmacı "Araştırma Uygulama İzinleri Genelgesine (2024/41)" göre belirtilen kapsamda araştırmasını yapmayı taahhüt etmiştir. Araştırmacının bilgi ve belgelerinin uygunluğu kontrol edilmiş olup aşağıda ifade edilen bilgiler kapsamında araştırma uygulama izni Millî Eğitim Bakanlığı ilgili birimleri tarafından onaylanmıştır.

Uygulama Yapılacak İller	Uygulama Yapılacak Birimler	Uygulama Yapılacak MEB Teşkilatları	Uygulama Yapılacak MEB Teşkilatının Kurum Kodu
KONYA	Anadolu Lisesi	[redacted]	325779

Not: *Okul/kurum yöneticileri tarafından "Araştırma Uygulama İzni" belgesinin ve veri toplama araçlarının (araçlardaki maddelerinin) modülde yer alan belge ve araçlarla aynı olduğu kontrol edilmelidir. Belgeler aynı olmadığı durumda araştırma uygulama izni verilmeyecektir.*

Doğrulama Kodu: 06fe2ddadd3d9709052148a34c265944e42df60eb9ddee8091be8ff09d95acff
Doğrulama Adresi: arastirmaizinleri.meb.gov.tr/belge-dogrula
[redacted]
e-posta: ttkb_arastirmaizinleri@meb.gov.tr, İnternet adresi: ttkb.meb.gov.tr

Sayfa 1/1

Appendix 13 Institute Board of Directors' Decision

Evrak Tarih ve Sayısı: 11.02.2025-E.535887

	T.C. TOKAT GAZİOSMANPAŞA ÜNİVERSİTESİ Kurul/Komisyon Kararları Formu	Doküman No	TOGÜ.FRM.185
		İlk Yayın Tarihi	02.08.2024
		Revizyon Tarihi	20.01.2025
		Revizyon No	01
		Sayfa No	1/1

Birim Adı	**LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**
Kurul/Komisyon Adı	**Enstitü Yönetim Kurulu**

Toplantı Tarihi	Oturum No	Karar No
08/02/2025	08	2025-08/ 01-172

Tokat Gaziosmanpaşa Üniversitesi Lisansüstü Eğitim Enstitüsü Enstitü Yönetim Kurulu, Enstitü Müdürü Doç. Dr. Ercan ÇAÇAN başkanlığında toplanarak aşağıdaki kararları almıştır.

KARAR NO 2025-08/137: Yabancı Diller Eğitimi Anabilim Dalı Başkanlığı'nın 03/02/2025 tarih ve 531479 sayılı yazısı ve ekleri okundu.

Buna göre; Yabancı Diller Eğitimi Anabilim Dalı, İngiliz Dili Eğitimi programı 239135656 numaralı Yüksek Lisans öğrencisi **Mehmet TAŞDEMİR** tarafından Dr. Öğr. Üyesi Zafer SUSOY danışmanlığında hazırlanmış olan tez konusu önerisinin "Yapay Zekâ Sohbet Botlarının EFL Lise Öğrencilerinin Konuşma Becerisi, Kaygı, Motivasyon ve Dijital Okuryazarlık Üzerindeki Etkisinin İncelenmesi" başlığıyla kabulüne ve 07/02/2025 (2024-2025 Bahar Yarıyılı) tarihi itibari ile tez dönemine geçmesine oy birliğiyle karar verildi.

Doç. Dr. Ercan ÇAÇAN
Enstitü Müdürü

Doç. Dr. Tuba KARABEY
Müdür Yardımcısı

Dr. Öğr. Üyesi Neslihan ÇIKRIKÇI
Müdür Yardımcısı

Prof. Dr. Özgür ALPARSLAN
ÜYE

Prof. Dr. İlhan EROĞLU
ÜYE

Prof. Dr. Sedat KARAMAN
ÜYE

Dr. Öğr. Üyesi Neslihan ÇIKRIKÇI
Enstitü Sekreteri V.
Raportör

Appendix 14 Ethics Committee Decision

	TOKAT GAZİOSMANPAŞA UNIVERSITY **Decisions of the Social and Human Sciences Research Ethics Committee Form**	Document No	TOGÜ.FRM.505
		First Publ. Date	19.02.2025
		Revision Date	
		Revision No	00
		Page No	1/1

Unit	General Secretariat
Committee	Social and Human Sciences Research Ethics Committee

Date of Meeting	Session Number	Decision Number
05.03.2025	04	01-72

The University Ethics Council for the Research in Social and Human Sceincies convened under the presidency of Prof. Dr. Tuğba KILIÇER.

DECISION 04.70- The act dated 10.02.2025 and numbered 534929 by The Directorate of Graduate School of Education was discussed.

It was unanimously decided that the applications to be performed and the data collection tools to be used by the reseachers whose details are given below conform to ethical principles.

THE TYPE OF THE RESEARCH STUDY	Master's Thesis
THE TITLE	An Investigation of AI Cahatbots on EFL High-School Learners' Speaking Proficiency, Anxiety, Motivation and Digital Literacy
THE THESIS ADVISOR/ AUTHOR	Assistant Prof. Dr. Zafer SUSOY Mehmet TAŞDEMİR (Department of Foreign Languages Education)
THE OPINION OF THE REPORTER	ADEQUATE

Prof. Dr. Tuğba KILIÇER
The Chairman of the Ethics Council
(Signature)

Prof. Dr. Mehmet Serkan UMUZDAŞ
Member
(Signature)

Prof. Dr. Emine ÖĞÜK
Member
(Signature)

Prof. Dr. Mehmet KARGÜN
Member
(Signature)

Associate Prof. Dr. Elif BOYRAZ
Member
(Signature)

Associate Prof. Dr. Ümran ÇEVİK GÜNER
Member
(Signature)

Associate Prof. Dr. Hüseyin Baha ÖZTUNÇ
Member
(Signature)