T.C

TOKAT GAZİOSMANPAŞA UNIVERSITY

POSTGRADUATE EDUCATION INSTITUTE

DEPARTMENT OF FOREIGN LANGUAGES EDUCATION

ENGLISH LANGUAGE EDUCATION MASTER'S DEGREE PROGRAM

ANXIETY LEVEL, EXAMINER'S GENDER, AND SPEAKING TOPICS: THEIR IMPACT ON EFL LEARNERS' PERFORMANCE

MASTER'S DEGREE THESIS

Bouchra BOURKHIS

Supervisor: Assist.Prof.Dr. Elham ZARFSAZ

TOKAT- 2025

ETHICAL AGREEMENT

According to the thesis writing guide of Tokat Gaziosmanpaşa University Graduate Education Institute, I declare that the Master's thesis titled “Anxiety Level, Examiner's Gender, and Speaking Topics: Their Impact on EFL Learners' Performance ”, which I prepared under the supervision of Assistant.Prof.Dr. Elham ZARFSAZ is an original work in accordance with scientific ethical values and rules, and I accept all legal sanctions if the contrary is detected.

…/…/…

Bouchra BOURKHIS

JURY APPROVAL

The defence exam of the thesis study titled “Anxiety Level, Examiner's Gender, and Speaking Topics: Their Impact on EFL Learners' Performance” prepared by Bouchra BOURKHIS was held on 03/09/2025 and has been accepted by the Jury below by Majority Vote, as a Master’s Thesis in English teaching as a Foreign Language at Tokat Gaziosmanpaşa University Graduate Education Institute English language teaching.

Members:	Signatures:
Member (Head of Committee): Dr. Öğr. Üyesi Şeyma KÖKCÜ	----------
Member (Supervisor): Dr. Öğr. Üyesi Elham ZARFSAZ	----------
Member (Committee Member): Dr. Öğr. Üyesi Zekiye ÖZER ALTINKAYA	----------

CONFIRMATION

...../...../.....

Director of Graduate Education Institute

ACKNOWLEDGMENT

I would like to express my deepest gratitude to my supervisor, Dr. Elham ZARFSAZ, who guided my work with insightful feedback and unwavering encouragement at all phases of this thesis. Her scholarship and dedication to intellectual excellence have not only informed the course of this study but also challenged me to think more profoundly and more clearly in my own work.

I would also like to thank the academic members and administrative staff of the Department of English Language Teaching at Gaziosmanpaşa University. The positive academic environment, accessibility of assistance, and their professionalism helped much in terms of the completion of this study.

Finally, I would like to offer my deepest gratitude to the respondents who participated in this study. Making their time and lives available in this work was only through the respondents' cooperation themselves, and I am indeed greatful for that.

ÖZET

KAYGI DÜZEYİ, SINAV GÖREVLİSİNİN CİNSİYETİ VE KONUŞMA KONULARININ YABANCI DİL ÖĞRENCİLERİNİN PERFORMANSI ÜZERİNDEKİ ETKİSİ

Bourkhis, Bouchra

Yüksek Lisans, İngiliz Dili ve Eğitimi Bilim Dalı

Tez Danışmanı: Doç. Dr. Elham Zarfsaz

Eylül 2025, xii + 70 sayfa

Bu çalışma, konu türü, sınavı yapanın cinsiyeti ve öğrenen kaygısı değişkenlerinin İngilizceyi yabancı dil olarak öğrenen (EFL) öğrencilerin konuşma performansı üzerindeki etkisini incelemektedir. Araştırmada öğrencilerin tarafsız konular ile cinsiyet önyargısı içeren konular üzerindeki performans farklılıkları, konuşma sınavları öncesi ve sonrası kaygı düzeylerindeki değişimler, ayrıca sınavı yapanın cinsiyetinin performans üzerindeki etkileri ele alınmıştır. Katılımcılar, B1–B2 düzeyinde İngilizce yeterliliğine sahip üniversite öğrencileridir ve farklı koşullar altında konuşma sınavlarına girmişlerdir. Veriler, Yabancı Dil Sınıf Kaygısı Ölçeği (FLCAS), yarı yapılandırılmış görüşmeler ve analitik puanlama rubrikleri aracılığıyla toplanmıştır.

Bulgular, öğrencilerin tarafsız konularda, cinsiyet önyargısı içeren konulara kıyasla daha yüksek performans sergilediklerini göstermiştir. Bu durum, konu aşinalığı ve tarafsızlığının sözlü değerlendirmelerde başarı için kritik olduğunu vurgulayan güncel çalışmaları desteklemektedir (Kim & Santos, 2023). Kaygı düzeylerinde sınav öncesinden sonrasına küçük bir azalma görülmüş, ancak bu fark istatistiksel olarak anlamlı bulunmamıştır; bu sonuç, EFL kaygısına ilişkin uzun dönemli araştırmalarla paralellik göstermektedir (Al-Mutairi & Nguyen, 2022). Ayrıca, kaygı ile konuşma performansı arasında negatif bir ilişki ortaya çıkmış, özellikle cinsiyet önyargılı konularda bu durum daha belirgin olmuştur (Rahimi & Zarei, 2024). Sınavı yapanın cinsiyetinin, özellikle konu türüyle birlikte değerlendirildiğinde, öğrencilerin performansını etkilediği görülmüştür; bu bulgu yüksek riskli sınavlarda giderek daha fazla tartışılan bir konudur (Chen & Ward, 2023). Nitel veriler, öğrencilerin kaygılarını azaltmak için

ön hazırlık, zihinsel prova ve gevşeme stratejilerini kullandıklarını ortaya koymuş, bu da alandaki pedagojik önerilerle uyum göstermiştir (Ibrahim & Lee, 2022).

Genel olarak bu çalışma, konuşma sınavlarında duygusal, bağlamsal ve kişilerarası etkenlerin etkileşimini ortaya koymaktadır. Bulgular, dil testlerinde duygusal boyutların önemine dikkat çekmekte ve EFL bağlamlarında daha adil ve kapsayıcı değerlendirme uygulamalarına duyulan ihtiyacı vurgulamaktadır.

Anahtar Kelimeler: EFL Öğrencileri, Konuşma Değerlendirmesi, Performans, Kaygı, Sınavı Yapanın Cinsiyeti, Konu Önyar

ABSTRACT

ANXIETY LEVEL, EXAMINER'S GENDER, AND SPEAKING TOPICS: THEIR IMPACT ON EFL LEARNERS' PERFORMANCE

Bourkhis, Bouchra

Master’s Thesis, English Language Teaching

Supervisor: Asst. Prof. Dr. Elham ZARFSAZ

September 2025, xiv+70 Pages

The current study explores the effect of topic type, examiner's gender, and learner anxiety on the speaking ability of English as a Foreign Language (EFL) learners. The study particularly focuses on finding out whether the learners differ in responding to neutral and gender-biased topics, how their anxiety levels change after and prior to speaking tests, and if the examiner's gender impacts the performance results. The participant was university-level learners at the B1–B2 levels of proficiency, who performed speaking tests under diversified conditions. The data were gathered with the use of the Foreign Language Classroom Anxiety Scale (FLCAS), semi-structured interviews, and analytical scoring rubrics.

The results found that the learners significantly outperformed on neutral topics compared to the gender-biased topics, substantiating recent findings that emphasize the importance of topic familiarity and neutrality in the speaking test (Kim & Santos, 2023). Whilst the levels of anxiety decreased slightly from the pre-test to the post-test, the change was statistically insignificant, consistent with long-term EFL anxiety research (Al-Mutairi & Nguyen, 2022). Anxiety was negatively correlated with speaking proficiency, including specifically in the tasks with the gender-biased topics, in agreement with evidence that anxiety compromises speaking (Rahimi & Zarei, 2024). Additionally, the gender of the examiner significantly impacted the end-state of the tasks when interacting with topic type, something more well-documented in the high-stakes test environment (Chen & Ward, 2023). The qualitative results also showed that the learners benefited from strategies such as preparing ahead, thinking they would say (cognitive rehearsal), and utilizing relaxation measures in order to cope with the experience of anxiety, in agreement with pedagogic guidance in the field (Ibrahim & Lee, 2022).

Overall, the research illuminates the interaction of emotional, contextual, and interpersonal factors in speaking tests. The study is among the growing body of research on affective effects in language testing and suggests more equitable, inclusive measurement processes in EFL contexts.

Keywords: EFL Learners, Speaking Assessment, Performance, Anxiety, Examiner Gender, Topic Bias

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

EFL: English as a Foreign Language
GB: Gender-Biased
GN: Gender-neutral
FLA: Foreign Language Anxiety
FLSA: Foreign Language Speaking Anxiety
EVT: Expectancy–Value Theory
FLCAS: Foreign Language Classroom Anxiety Scale
FNE: Fear of Negative Evaluation

CHAPTER I

INTRODUCTION

Background of the Study

Language learning, particularly in foreign or second language contexts, is not merely a matter of acquiring linguistic forms but also of navigating a complex interplay of cognitive, affective, and social dimensions. Among these, speaking anxiety has been consistently identified as one of the most influential affective variables shaping learners' performance and overall success in acquiring communicative competence. Foreign Language Anxiety (FLA), as conceptualized by Horwitz, Horwitz, and Cope (1986), refers to a distinct form of anxiety that arises specifically in the context of language learning, manifesting through communication apprehension, test anxiety, and fear of negative evaluation. While anxiety has been widely studied across learning domains, it is particularly debilitating in oral communication, where learners must process meaning in real time, retrieve appropriate linguistic forms under pressure, and negotiate meaning with interlocutors. Such demands intensify cognitive and emotional strain, often leading to disrupted fluency, avoidance behaviors, and diminished performance (Horwitz, 2017; MacIntyre, 2017).

Anxiety interacts with multiple variables in speaking contexts, ranging from learners' linguistic proficiency to external factors such as topic familiarity, task type, and interlocutor characteristics. The willingness to communicate, as highlighted by MacIntyre, Clément, Dörnyei, and Noels (1998), is closely tied to learners' perceptions of their linguistic resources and the social conditions under which communication occurs. Research has shown that when learners are asked to discuss unfamiliar or abstract topics, the cognitive load placed on working memory is heightened, leaving fewer resources available for fluent and accurate language production (Chandler & Sweller, 1991; Sweller, 1988). Conversely, familiar or personally relevant topics can reduce this strain, enabling learners to allocate more resources to meaning-making and interaction (Sweller, Ayres, & Kalyuga, 2011). These insights, drawn from Cognitive Load Theory, suggest that task design in oral assessments should carefully consider both linguistic and cognitive demands if learners are to demonstrate their actual communicative abilities rather than their ability to cope with stress Sweller, Roussel, & Tricot, 2022).

Test examiners' function is another important variable that defines speaking performance. Long's (1996) Interaction Hypothesis identifies the enabling function of conversational interaction in second language acquisition as assuming that negotiation of meaning, clarification requests, and modified input produce more intelligible and accessible communication. The interaction pattern of the examiner in an oral test, however, may vary with variables such as personality, culture, and gender. Examiner gender, in this context, has proven to be a subtle but significant variable influencing learners' comfort level, anxiety, and interaction opportunities (Pica, 1987; Wei, 2012). Students can vary in how they behave in the presence of male and female examiners, with diverse effects in how they take risk, respond in answering questions, or handle breakdowns in communication. What this suggests is that the examiner's gender is not some other demographic variable but something that may condition the dynamics of the oral test-taking experience.

Emotional and affective factors are significant in the manner in which people learn and use the second language. Krashen's (1985) Affective Filter Hypothesis is the manner in which anxiety, weak motivation, and lack of confidence can be an emotional block that prevents the learner from adopting and using the language in an effective manner. If the block is high, learners are unable to draw on their full language ability, even if they understand the material. If the block is low, they are more effectively using the language. There is evidence in favor of this idea: Hu (2023) demonstrated that anxiety was capable of significantly impacting the speaking English ability of learners, while being motivated and feeling confident was reported by Du (2009) as decreasing this barrier. Anxiety at being assessed was also found by Ren et al. (2025) as negatively impacting speaking ability when the learner lacked the use of self-assessment tools. What this suggests is that how speaking tasks are devised and who is marking are crucial in speaking tests. Interesting or familiar material and markers who are encouraging and supportive will decrease the emotional block so that the learner is better at using their language ability. If the material is more than they can manage or the markers are stern, the block is more intense so that the learner is unable to function. Thus, understanding of these emotional factors is significant in terms of understanding how well the learner does in speaking tests.

Against this background, Task-Based Language Teaching (TBLT) provides an insightful pedagogic account of the relationship between tasks, topics, and learner interest.

Willis and Willis (2007) provide a three-stage model—pre-task, task, and post-task—that integrates fluency with accuracy with a focus on authentic use.

Under this condition, selecting the appropriate topic is no trivial issue; it contributes significantly towards how much learners are motivated, how anxious they become, and how effectively they perform. If topics are closely related to what learners know, care, and are prepared to ponder, they will be in the position of making learning more interesting, decreasing stress, and enabling the students to meaningfully communicate (Krashen, 1986; Schneider et al., 2018). If topics fail to suit or rely on other cultures, they will cause learners anxiety, constrain the extent to which they will be in the position to fluently speak, and will be unable to unveil the extent of what they actually know (Alazemi, 2023; Gkintoni, 2025). The current study explores the way in which they intersect, acknowledging that the extent to which a topic causes the learner to reflect, the nature of the interpersonal relationship, the role of affect in learning, and the way in which the instruction tasks are conducted all contribute toward the way in which the students perform in speaking (Krashen, 1986; Gkintoni, 2025).

Statement of the Problem

Speaking proficiency is one of the most problematic EFL learning skills, particularly in high-stakes speaking tests. Research always finds that proficiency is shaped not only by linguistic proficiency but also by psychological as well as contextual variables such as test anxiety, traits of the examiner, as well as speaking topic type (Zhang & Wang, 2022; Kiruthiga & Christopher, 2022). Often, students get more anxious in the face of oral exams, suppressing fluency as well as accuracy, resulting in scores which may be other than what they are capable of in the language (Chen & Li, 2023).

Speaking topic selection is another issue. Gender-biased, culturally unfamiliar, or unfamiliar topics may induce high cognitive load and high anxiety, thus reducing the efficiency of learners' speaking production (Kim & Park, 2023; Loveless, 2023). Gender-neutral or familiar topics decrease affective barriers and stimulate more active use and more correct use. Though this is the case, little research systematically compared the influence of topic type on learners' speaking proficiency under the condition that the mediatorial role of anxiety is controlled.

Gender of the examiner is another subtle but crucial variable that underlies the interaction dynamics in oral exams. Some pieces suggest that students will perform better under same-sex examiners because they will be more at ease and experience more empathetic feelings (Rahimi & Alavi, 2023), whereas others assert the opposite (Rahimi & Alavi, 2023). Nevertheless, the majority of the related work was conducted based on Western cultures, with sparse empirical support from EFL contexts, so there are questions surrounding cultural and contextual generalizability (Zhang & Liu, 2023).

Against this context of lacunae, there is a pressing research necessity in investigating the interaction of topic type (gender-biased vs. gender-neutral), learner anxiety, and examiner gender in influencing EFL learners' speaking skill. The present study bridges this gap by investigating their simultaneous effect in the university EFL context, and providing insights with practical as well as theoretic applications in the conceptualization of tests.

Significance of the Study

The value of this research is that it may add both theoretically and practically to the research in language assessment. At the level of theory, the research draws upon cognitive, affective, and interactionist perspectives in order to emphasize the multivariable nature of speaking performance as an effect produced through the interaction of various interacting variables. Through the study of topic type and examiner gender as well as that of learner anxiety, the research makes the more comprehensive account of the contexts of oral assessment.

In practice, the results have important practical implications for teachers, examiners, and test designers. For teachers, the improved understanding of how the selection of topic and examiner demeanor influence learner anxiety can inform classroom practices that foster emotional safety and equitable engagement. For examiners, the research draws attention to the importance of being aware and sensitive to their interactional style as well as the potential impact on the performance of the learners. For test designers and institutions, the research draws attention to the importance of ensuring that the practices of assessment be equitable, culturally sensitive, and in accord with the cognitive and affective realities of the learners. Overall, the research endeavours to inform the development of speaking exams that more accurately reflect learners' communicative competence with little avoidable bias or undue anxiety.

Objectives of the Study

Four general goals guide the study in the following. First, the study seeks to investigate the relationship between foreign language anxiety and EFL learners' speaking ability in speaking tasks. Second, the study tries to examine the influence of the examiner's sex as influencing the extent of learners' anxiety and speaking ability. Third, the study seeks to investigate the influence of topic type in speaking with both neutral and biased topics in determining the learners' performance under test situations. Fourth, the study seeks to investigate the independent and interacting effects of the variables, demonstrating an in-depth representation of how anxiety, the examiner's sex, and topic selection interact in the context of the oral examination.

Delimitations and Limitations

Like any study, the present study is bounded within some delimitations and limitations. The study is delimitated to the population of EFL learners at the intermediate proficiency level, and the findings may not be generalized across all levels of proficiency or situations. The study is limited in its examination of speaking tests within one institutional test condition type, which may differ in design and stakes from other test situations. Additionally, while the study is limited in the study of the external variable examiner gender and topic type, the study is not comprehensive in including other potential influencing variables such as examiner training, culture, or learners' broader motivational profiles. These delimitations represent required constraints within which to conduct an in-depth study, while the limitations indicate the interpretive circumspection required in generalizing findings from the immediate research context.

Research questions

Drawing on the conceptual framework, which highlights the interaction between anxiety, speaking topics, and examiner gender, this study is guided by the following research questions:

1. Is there a significant difference in EFL learners’ speaking performance between gender-neutral and gender-biased speaking topics?

2. Do learners' anxiety levels change significantly from pre-test to post-test after the speaking tasks?
3. Is there a significant relationship between learners' anxiety and their speaking performance scores?
4. Does examiner gender significantly affect learners' performance differences while controlling for anxiety?

Definition of Key Terms

- **Gender-biased topics:** Gender-biased topics involve preferential treatment, stereotypes, and discrimination based on gender identity, often favouring one gender over others, typically men. This bias can manifest in various areas like the workplace, education, media, and social interactions, impacting non-binary and transgender individuals differently (IMD Business School, 2025; Drew, 2023).
- **Gender-neutral topics:** Gender-neutral topics utilise inclusive language and concepts to prevent gender bias, promoting respect and equality for all gender identities, a crucial approach for creating inclusive and respectful communication and environments (M. M. MSEd, 2021; Hutto, 2024).
- **Examiner gender:** The "gender of the examiner" refers to the gender the examiner identifies as, which can impact the interaction dynamics between the examiner and the test-taker or candidate during oral or speaking evaluations **(**Wood,n.d).
- **Foreign Language Anxiety (FLA) / Speaking Anxiety** In this study, FLA refers specifically to *speaking-related anxiety* experienced by learners during oral performance tasks in English, characterized by feelings of nervousness, apprehension, or fear of negative evaluation. It is measured using a modified version of the [name of scale you use, e.g., Foreign Language Classroom Anxiety Scale – Horwitz et al., 1986] adapted to focus on speaking tasks. This operational definition emphasizes the *situational* nature of anxiety, occurring immediately before, during, and after the speaking assessment.
- **Intersectionality:** is a theory that explains how diverse social and political identities, such as gender, caste, sex, race, ethnicity, class, sexuality, religion, disability, physical appearance, and age, create distinct forms of discrimination and privilege (Deckha, 2008)

- **Oral performance:** Oral performance art, patterned performative speech for an audience, is perhaps the oldest and most ubiquitous human art form (Ugoretz, n.d.-b).

CHAPTER II

LITERATURE REVIEW

Introduction

Speaking is the foundation of efficient communication in order that individuals would be in a position to communicate effectively, with reduced ambiguity and misinterpretation (Sharma, 2024). The speaking skill is instrumental at both the personal and professional levels through the attainment of self-confidence, cultivating self-assertion as well as clearing the way through social as well as professional development (Qureshi, 2020). When we venture into the field of learning foreign languages, speaking continues to be the most challenging skill to learn due to cognitive, affective, as well as social impacts on the performance.

The following chapter is a review of works on three related aspects core to the current study: (1) anxiety, with special reference to foreign language speaking anxiety (FLSA), (2) topic of speaking and its cognitive demand, and (3) sex of the examiner and its effect on communication and performance. The account weaves together definitions, argument from theory, evidence from research with focal emphasis on conceptual contrast as well as applicability within study. The chapter is summed up with suggestion of lacunae in the published work that is being filled in this current study.

Understanding Anxiety

General Concept of Anxiety

Anxiety is broadly described as an unpleasurable affective state accompanied by feelings of fear, apprehension, or anxiety regarding future occurrences (Castelfranchi, 2014; Crocq, 2015). While fear reacts to a present danger, anxiety is a projection of fear regarding potential future harm. It becomes generalized and unfocused and often manifests as an exaggeration about situations subjectively found to be threatening (Bouras, 2017).

Varied authoritative sources converge to agree that anxiety involves psychological as well as physiological components. Many authoritative sources concur that anxiety involves psychological as well as physiological components. The American Psychological Association records symptoms such as muscle tension, intrusive thoughts, and a increased heart rate. While the Encyclopedia Britannica highlights its ability to impair self-esteem and

trigger avoidance behavior, UNICEF cites anxiety as a normal human response to perceived threats that typically resolves once threats are removed. O'Neill and Sorochan (2014) describe its chronic form as a lingering negative affect accompanied by heightened alertness and stress-related physiological response. Understanding its general properties grounds a focus on its particular expression within speaking and foreign language contexts. FLSA broadens the broader anxiety construct but is guided by the exceptional pressures of language performance.

Speaking Anxiety

Speaking anxiety, often referred to as *glossophobia*, is the fear or apprehension experienced before or during oral communication, particularly in public contexts (Hancock et al., 2009). It may present through physical symptoms (e.g., sweating palms, shaky voice), cognitive symptoms (e.g., self-doubt, fear of failure), and behavioral symptoms (e.g., avoidance of speaking situations) (Cunill, n.d.; Montijo, 2022).

In language learning contexts, Gardner and MacIntyre (2013) define speaking anxiety as the apprehension that arises when learners are required to perform in a second or foreign language. Brown (2000) situates it as an affective factor directly influencing oral performance, while Zhang (2011) emphasizes its physical and psychological manifestations during language tasks. Samuelsson (2011) frames it as a specific social phobia affecting up to 20% of the population, underscoring the prevalence of this phenomenon.

Horwitz et al. (1986) identify speaking anxiety as a form of communication apprehension, driven by concern over others' evaluations and characterized by avoidance of evaluative situations. Across definitions, the key dimensions include anticipation of negative evaluation, perceived inadequacy in performance, and resulting physical arousal. Speaking anxiety is the primary affective variable under investigation, making it critical to distinguish its features from other forms of communication apprehension. This clarity informs the theoretical framing and interpretation of the study's findings.

Foreign Language Speaking Anxiety (FLSA)

FLSA or Foreign Language Speaking Anxiety refers to a special kind of foreign language anxiety--also known as xenoglossophobia--characterized by nervousness, anxiety, and discomfort related to speaking (in) a foreign language (MacIntyre & Gardner, 1994).

Horwitz et al. (1986) describe foreign language anxiety as a "distinct complex of self-perceptions, beliefs, feelings, and behaviors found as a set within the classroom language learning situation." FLSA is limited to speaking production only.

Camden (2020) describes it as tension and nervousness triggered by the autonomic nervous system in foreign language speaking contexts. Öztürk (2012) further specifies that FLSA can disrupt learning processes, influenced by gender, proficiency level, and motivation. Given that the research examines speaking anxiety in EFL learners, FLSA represents the precise operational construct of interest, providing a targeted lens through which the impact of examiner gender and topic type can be assessed.

Table 1. Summary of Definitions: Anxiety, Speaking Anxiety, Foreign Language Speaking Anxiety

Concept	Definition	Key Sources
Anxiety	An unpleasant emotional state characterized by apprehension and worry about potential future events; it differs from fear, which is a response to an immediate threat.	Castelfranchi (2014); Crocq (2015); Bouras (2017)
Speaking Anxiety	Nervousness and apprehension occurring before or during speaking tasks; includes physical, cognitive, and behavioral symptoms.	Hancock et al. (2009); Gardner & MacIntyre (2013); Horwitz et al. (1986)
Foreign Language Speaking Anxiety (FLSA)	A type of foreign language anxiety specific to oral performance in a second/foreign language, often marked by nervousness, tension, and fear of negative evaluation.	MacIntyre & Gardner (1994); Horwitz et al. (1986); Öztürk (2012)

Theoretical Perspectives on Language Anxiety

A number of theoretical models have been proposed to account for the role of emotional, cognitive, and contextual variables on speaking performance and language learning. The present section discusses those most pertinent to this study with an emphasis on their central assumptions and foreign language speaking anxiety (FLSA) related findings.

Horwitz's Theory of Foreign Language Anxiety

Horwitz, Horwitz, and Cope (1986)'s conceptualization of Foreign Language Anxiety (FLA) remains one of the most influential applied linguistics frameworks to date. They define FLA as "a distinct complex of self-perceptions, beliefs, feelings, and behaviors related to classroom language learning arising from the uniqueness of the language learning process" (p. 128). Theoretically speaking, within this model, there are three related constituent components which affect learners' emotional experience:

a) Communication Apprehension: Apprehension about communication is anxiety or fear related to actual or future communication with other people (McCroskey, 2001). The anxiety usually comes with physiological symptoms such as pounding heart, shaking hands or dizziness (Sapolsky, 2004). For language learners, this anxiety is exacerbated by limited control over language and perceived observation by one's peers and instructor (Horwitz et al., 1991).

b) Fear of Negative Evaluation (FNE): FNE entails anticipation that others will negatively judge one's performance (Yılmaz & De Jong, 2024). It is especially prominent in speaking tasks on oral language proficiency tests, as errors regarding pronunciation or grammar are liable to be interpreted as signs of incompetence (Aydın, 2008). Students with perfectionist feelings or low self-esteem are especially susceptible to FNE, which lowers their desires to engage with speaking exercises (Bozdağ, 2021; Van Der Molen et al., 2014).

c) Test Anxiety: Test anxiety is performance anxiety that is associated with fear of failure during evaluative situations (Howe, 2014). It can lead to disruptions in thinking (e.g., forgetting) and physiological arousal during oral exams, inhibiting speech production (Laurin-Barantke et al., 2016; Zambak, 2024).).

Relevance to the current study: This model applies directly because it simultaneously encompasses performance anxiety as well as language-related difficulties. It gives a holistic perspective on understanding how speaking topics and examiner gender might intersect with anxiety within speaking test contexts.

Affective Filter Hypothesis (Krashen, 1982)

Krashen's Affective Filter Hypothesis suggests that affective variables—chiefly motivation, self-confidence, and anxiety—act as a "filter" to either enable or impede language input processing. If the affective filter is high, as with high anxiety or low motivation, efficiency of language acquisition is lowered; with a low filter, more successful engagement with input is enabled (Rost, 2022). Motivation is key, with highly motivated learners actively involved with tasks and successful at language acquisition (Liu, 2023; Yaoqing, 2021). Self-confidence also contributes to learning through fostering communicative risk-taking as well as decreasing fear of negative evaluation (Getie, 2020). Anxiety puts the filter in place, which limits exposure to manageable input as well as inhibiting performance (Kiruthiga & Christopher, 2022). Krashen's account reflects how affective variables meet environmental factors like topic knowledge and examiner attitude at the center of comprehending speaking anxiety dynamics as well as speaking test in EFL contexts.

Interaction Hypothesis (Long, 1983)

Long's Interaction Hypothesis highlights the focal status of conversational interaction in second language acquisition (SLA), with language acquisition deriving advantage from the receipt of comprehensible input as well as engagement in meaning negotiation. The latter usually entails modifed interaction (Long, 1996). Negotiation of meaning requires strategy types such as clarification request, confirmation check, or reformulation in order to reduce breakdowns in communication and arrive at comprehension (Pica, 1987). Therefore, modifed interaction—the slower speech rate, the use of reduced vocab, or repletion—renders the input more transparent, particularly native-non-native speaker interactions (Wei, 2012). The interactive style of the examiner in the context of oral exams may have dramatically influential bearings upon the quality of input and opportunities for negotiation and upon the learners' level of fluency, level of confidence, and level of performance.

Expectancy-Value Theory (Eccles et al., 1983)

Expectancy–Value Theory (EVT) proposes that learners' motivation depends on their expectation of success and the value they attach to the task. Task value is further divided into intrinsic value (enjoyment), utility value (usefulness), attainment value (personal

importance), and cost (effort, anxiety, or loss of alternatives) (Eccles & Wigfield, 2002). Speaking topics perceived as interesting, useful, or personally relevant are likely to lower anxiety and increase engagement. In contrast, unfamiliar or low-value topics may heighten stress, especially under the pressure of oral assessment.

Table 2. Summary of Key Theories Related to Language Anxiety — Concepts, Main Proponents, and Relevance to Study

Theory	Key Proponents	Core Principles	Relevance to Current Study
Horwitz's Theory of FLA	Horwitz, Horwitz & Cope (1986)	Anxiety in language learning arises from communication apprehension, fear of negative evaluation, and test anxiety.	Provides a comprehensive framework for understanding FLSA and its performance effects.
Affective Filter Hypothesis	Krashen (1982)	Emotional states (motivation, self-confidence, anxiety) act as a filter that facilitates or blocks language input processing.	Highlights how examiner demeanour and topic familiarity may influence learner receptivity and output.
Interaction Hypothesis	Long (1983, 1996)	Comprehensible input and negotiation of meaning during interaction promote SLA.	Suggests that the examiner–candidate interaction style may impact speaking performance.
Expectancy–Value Theory	Eccles et al. (1983)	Motivation depends on the expectation of success and task value (intrinsic, utility, attainment, cost).	Explains how topic choice affects learner engagement and anxiety.

Speaking Topics and Language Performance

Speaking topics are what or which subject or theme students are required to speak about during speaking exercises. They affect what level of comfort, encouragement, and success students are during speaking exercises. This depends on whether they are comfortable with or sensitive to a subject (Bachman & Palmer, 1996; Fulcher, 2003). Language performance is when an individual uses language adequately well in true communicating situations. It shows how fluent they are speaking, what level of grammar they use accurately, what level of organisation of thoughts they have, and what wide vocab they have (Chomsky, 1965; Luoma, 2004). Research found that how well-versed a person is with a subject, what type of task they are doing, and even what an examiner is doing all can

affect speaking performance (Hughes, 2003; Brown, 2007). This shows that speaking topics and language performance are highly related to each other regarding speaking tests (Hughes, 2003; Brown, 2007). Individuals' speaking topics have a large effect on whether they speak well or poorly during speaking exercises. They affect both what mental processing is needed to do a job (Ellis, 2005; Robinson, 2001; Skehan, 1998) and what they are experiencing emotionally, such as anxiety or pleasure (Aubrey, 2017; Lambert & Zhang, 2019; Mora, Mora-Plaza & Bermejo Miranda, 2023). Research shows that if participants possess much knowledge about a subject, this aids speaking more fluently, accurately, and with more complex language because there is reduced mental pressure (Qiu, 2019). But if a topic is unfamiliar or uncomfortable to speak about, this might add to pressure and make performance harder (Nakamura et al., 2021). So, how complex a topic is, how the task is designed, and how familiar the learner is with the topic all strongly connect with their emotions and how well they can speak.

Cognitive Load Theory (Sweller, 1988)

The Cognitive Load Theory (CLT) states that working memory is of restricted capacity, and performance and learning are impaired when cognitive demand outweighs this capacity (Sweller, 1988). Cognitive load during speaking exercises derives from content difficulty, need for linguistic structure, and immediacy of on-line processing. Three categories of cognitive load are typically distinguished: intrinsic load that captures the essence of intrinsic difficulty arising from the number and interaction of items (Chandler & Sweller, 1991); extraneous load that derives from poor design of exercises or irrelevant distracters (Loveless, 2023); and germane load that represents mental effort invested in processing, constructing, and automating schemas for eventual acquisition (Sweller, 2010). Abstract or unfamiliar speaking subjects can elevate intrinsic load and thereby anxiety. Relatable or self-focused subjects reduce cognitive tension and free working memory resources for more effective language production.

Task-Based Language Teaching (Ellis, 2003)

Task-Based Language Teaching (TBLT) emphasizes using meaningful, real-life communicative tasks rather than decontextualized grammar or vocabulary drills. Its three-phase framework, *pre-task, task, and post-task*, guides learners through preparation, performance, and reflection (Willis & Willis, 2007). Recent empirical findings show that when task topics are personally meaningful, learners achieve higher fluency and report lower

anxiety (Kim & Park, 2023). In the pre-task phase, learners think about what they already know and get ready by understanding what they need to do in the activity. This helps them use their mind more easily (Ellis, 2003; Willis, 1996). The task phase centers on authentic interaction, where fluency is prioritized and the teacher assumes the role of facilitator rather than knowledge transmitter(Skehan, 1998; Ellis, 2009). Finally, the post-task phase shifts the focus to accuracy, feedback, and consolidation, enabling learners to refine their performance and internalize key language features(Willis & Willis, 2007).

When topics are important to learners and are part of a task-based learning approach, they can make speaking more interesting, less stressful, and help learners improve both how smoothly they speak and how correct their language is (Ellis, 2017; Lambert & Zhang, 2019; Aubrey, 2017).

Social Cognitive Theory (SCT) highlights the role of observational learning, self-efficacy, and self-regulation in skill acquisition. Learners who observe peers successfully performing speaking tasks and believe in their ability to do the same are more likely to participate confidently. Familiar topics can boost self-efficacy, while exposure to positive role models during practice may reduce the anxiety associated with oral tasks.

Impact on Fluency, Accuracy, and Vocabulary Use

Research shows that speaking anxiety can undermine fluency by disrupting speech flow, causing hesitations, and limiting lexical variety (Kormos & Dörnyei, 2004; Donate, 2022). Conversely, when learners value familiar topics, they tend to produce more coherent, fluent, and lexically rich speech (Tennant, n.d.). Accuracy may be less directly affected by topic choice, but reduced anxiety can indirectly improve grammatical control.

Anxiety Modulation through Topic Selection

Abstract or unfamiliar topics require more cognitive effort, intensifying anxiety, especially in evaluative contexts (Montijo, 2022b). In contrast, concrete and personally relevant topics often reduce cognitive strain and create a sense of psychological safety, enabling learners to focus on communication rather than self-monitoring.

Table 3. Cognitive and Affective Implications of Topic Complexity in Speaking Tasks

Topic Type	Cognitive Load Impact	Affective Impact	Likely Performance Effect
Familiar, concrete topics	Lower intrinsic load; easier retrieval of vocabulary and structures.	Lower anxiety; higher self-efficacy.	Greater fluency, more coherent responses.
Abstract, unfamiliar topics	Higher intrinsic load; greater processing demands.	Increased anxiety; risk of cognitive overload.	Hesitations, reduced lexical variety, and possible accuracy decline.
Moderately challenging topics	Balanced cognitive demands; manageable load.	Optimal engagement; reduced boredom.	Potential for growth in fluency and complexity.

Examiner Gender and Communication

Examiner characteristics, particularly gender, can shape oral assessment interactions through verbal and nonverbal behaviors and test takers' perceptions and comfort levels.

Gender Schema Theory (Bem, 1981)

Sandra Bem's gender schema theory explains children learning sex-defined characteristics such as preference, skills, types of personality, behaviors, and self-concepts attached to gender. It attempts to show how gender-related representations within the mind come to fruition at young ages and account for cultural explanations of attention, drive, perception, impression formation, and action (Canevello, 2020). Gender schema theory assumes gender typing is derived from gender-related schematic processing by which individuals process information on gender-related connectivity within the gender schema. Assimilation of self-concept forms this action within the gender schema. Gender-typed individuals prove more liable to process on the basis of the gender schema, perhaps due to the high value put on functionally by culture on the gender dichotomy. Political uses of theory as well as its relation to androgyny are discussed too (Bem, 1981). When a student's gender schema associates authority or appraisal with a gender, this might change their level of anxiety as well as their communicative performance with examiners who are of this gender.

Empirical Evidence on Examiner Gender Effects

Research on examiner gender as part of the environment in the oral test finds variable but compelling patterns. Some students disclose enhanced performance with same-sex examiners, relation often summarized as due to the high level of comfort and perceived understanding (Buckingham, 1997). Others disclose enhanced response with opposite-sex examiners, as perhaps an artifact of cultural dictates or individual penchant (Porter, 1991a; O'Loughlin, 2002). At the more general level than variation at the pupil level come variations in "interactional style, often coordinated with norms for gendered communication," that influence the quantity and quality of linguistic production from interviewers in the context of the oral test.

Gendered Communication Styles

It is possible that women are more inclined to employ rapport-establishing and supporting conversation strategies and that men are more inclined to direct or competitive approaches (Coates, 2015; Geller, 2025). Nonverbal communication,—eye contact, nodding, and body positioning,—is also different and can affect candidate comfort levels. Differences in style are most salient when paired with speaking topic difficulty because they can either enhance or reduce anxiety.

Table 4. Summary of Empirical Studies on Examiner Gender and Speaking Performance

Author(s) & Year	Participants	Context	Main Findings
Porter (1991)	Mixed-gender EFL learners	Oral interviews	Higher scores with male examiners; cultural factors likely.
Buckingham (1997)	32 Japanese learners	Oral proficiency interviews	Interaction effect: females scored higher with female examiners; males with male examiners.
O'Loughlin (2002)	IELTS candidates	Oral test interviews	Examiner gender and style may influence both output and ratings.
Erdiana et al. (2019)	University EFL students	Speaking tests	Women showed higher fluency and sensitivity in speech.

Intersectional Perspective

Whereas previous sections explored anxiety, topic speaking, and gender of examiner individually, they tend to combine within oral assessment settings in complex and mutually supporting manners. Intersectionality as developed by Crenshaw (1989) is a conceptual tool to understand how more than one social and contextual variable come together to affect student experience. For speaking performance in oral assessment settings, this entails that performance is not determined by discrete variables but by the blending together of personal identity, task design, and interlocutor properties.

For example, a female EFL learner with high speaking anxiety might perform differently when discussing an unfamiliar, abstract topic with a male examiner than when discussing a familiar, personally relevant topic with a female examiner. Each variable—anxiety level, topic type, and examiner gender—could independently affect her performance, but their combined effect may be greater than the sum of their parts.

This integrated perspective highlights that oral assessments are socially situated events rather than purely linguistic evaluations (Menard-Warwick, 2008). Learners navigate the linguistic demands of speaking and the social dynamics of power, identity, and expectation. Consequently, understanding the intersections among affective, cognitive, and interpersonal factors is crucial for designing fair and valid assessment practices.

Related Empirical Studies

Several studies have examined the impact of anxiety level, examiner gender, and speaking topic on EFL learners' speaking performance, with particular attention to gender differences, the role of anxiety in speaking proficiency, and the influence of topic content on performance.

Patel and Kar (2025) investigated anxiety levels among first-year undergraduate students delivering English speeches in an EFL context. The study involved 164 students (78 females and 86 males) and employed Horwitz et al.'s *Foreign Language Classroom Anxiety Scale* (FLCAS) to measure anxiety. Findings indicated that female students reported higher levels of anxiety than their male peers. Statistical coding techniques revealed a significant relationship between public-speaking anxiety in English and gender. The internal

reliability of the Likert scales was confirmed using Cronbach's alpha. The authors argued that addressing gender variations in language anxiety is important for helping students overcome public speaking difficulties and improve academic performance and communication skills.

Similarly, Okyar (2023) examined the relationship between EFL speaking anxiety and self-efficacy in 293 English preparatory school students from a central Turkish state university. Using Öztürk's five-point Likert scale to measure speaking anxiety, the study found a negative correlation between anxiety and self-efficacy: students with higher self-efficacy reported lower anxiety levels. Data were collected via questionnaires administered to 11 classes during the fall semester of 2019–2020. The findings suggest that learners' belief in their ability to communicate effectively in English can reduce speaking anxiety. Consistent with Patel and Kar (2025), this study also reported that female students exhibited higher anxiety levels than males, although it noted that the strong influence of self-efficacy on emotional responses may explain variations in results compared to earlier research.

Rafek et al. (2014) also applied one version of the FLCAS as part of exploring levels of anxiety in second language acquisition. The findings indicated significant levels of anxiety in both male and female learners, with females again having the highest. The authors pointed out that stress from the activities of the language class, including speaking, listening, writing, and reading, may produce apprehension and negative associations if speaking is among the tasks required in the assessment. If there is no guidance in place, the activities would quickly become anxiety-provoking ones, with learners being at loss and disengaged.

Mohtasham and Farnia (2017) in another study explored the influence of gender on English-speaking anxiety among university EFL learners in Isfahan Province. Anxiety levels, reactions when confronted with various speaking tasks, causes, and alleviating measures were assessed. The study indicated that spontaneous speaking tasks contributed to both males' and females' anxiety. Females reported medium level of worry in general, whereas males were most perturbed when required to talk on interesting issues or when they were unprepared. The authors emphasized pedagogic significance in the identification of anxious learners and the implementation of measures reducing the distress in immediate speaking tasks in the classroom.

About examiner gender, Zulkiflee and Nimehchisalem (2022) conducted a mixed-methods study during the COVID-19 pandemic, using the *English as a Second Language*

Speaking Anxiety Scale (ESLSAS) and virtual semi-structured interviews. The sample comprised students from both single-gender and co-educational schools in Terengganu. The study addressed two questions: whether speaking anxiety levels differed between school types and between male and female students, and whether these differences were statistically significant.

Altakhaineh (2019) examined the influence of gender and culture on Arab IELTS candidates' speaking performance in the UAE, focusing on Al Ain. Based on interviews with male and female examinees and teachers, the study found that examiner gender could influence linguistic behaviour and performance. Many participants reported feeling hesitation, embarrassment, or discomfort when examined by someone of a different gender, which affected their responses. This finding contrasted with O'Loughlin's earlier work, which found no such effect.

The effect of speaking topic type—more specifically, gender-biased versus gender-neutral topics—has also been researched. Rojas-Buñay and Guaman-Luna (2024) indicated that topic relevance, familiarity, and type have significant effects on EFL learners' use of vocabulary, pronunciation, and interactional skill. Exciting, gender-neutral topics like social media, technology, and pop culture were linked with strong performance. The study by Kutuk et al. (2021) concluded that male learners' self-efficacy would be lowered when they were taught with gender-biased topics, causing effort, persistence, and achievement levels to be lowered. However, this may increase self-efficacy among female learners if the topics conformed to dominant gender expectations. The authors believed that using gender-neutral topics can prevent stereotype reinforcement, therefore aiding learners' confidence and interest regardless of the sex. Likewise, Liu et al. (2022b) and Azkarai (2015) found that using gender-neutral topics mitigates stereotype threat and cultivates equal participation, so speaking performance is enhanced with more equal classroom interaction.

Intersectionality (Crenshaw, 1989) is an insightful framework that sheds light on the role of intersecting social identities (e.g., gender, race, class, and language) in shaping learners' experience in the world of learning (Norton, 2013; Block, 2014). The identity of the examiner, topic content, and the anxiety in the learner intersect in a socially contexted world in speaking tests, influencing levels of comfort as well as linguistic output. Previous work in examiner gender (O'Sullivan, 2002) and speaking anxiety (Horwitz et al., 1986) often considered the two in relative independence. Intersectionality, on the other hand, recognizes in how many ways they may intersect in complex ways: that is, the same learner

may be different in response in responding to the same gender-biased topic under male vs. female examiner, especially in cultural contexts with describable, rigidly fixed gender-related conducts. Such entwinings have the effect that speaking proficiency indexes linguistic ability as well as the negotiation of identity, dynamics of power, as well as affective protection (Menard-Warwick, 2008; Motha, 2014).

Through an exploration of the interaction of examiner sex, learner anxiety, and type of speaking topic, the current study contributes to this new scholarship. Taking an intersectional stance enables us better to account for variability in performance as well as to respond more effectively to the challenge of more just, socially responsible forms of assessment.

Summary and Research Gaps

The study identifies some significant insights that directly inform the present study. First, Foreign Language Speaking Anxiety (FLSA) is introduced as an independent construct derived from general communication apprehension and language-related pressures (Horwitz et al., 1986). The Affective Filter Hypothesis, the Interaction Hypothesis, and Expectancy–Value Theory as classic theories bring out the interaction of affective states, the structure of the task, and learning attainment. Under the umbrella of such interaction, speaking topics take center stage in the development of cognitive overload and anxiety: speaking with familiar or personal topics would be linked with fluency and confidence. However, speaking with unfamiliar or sensitive topics would create stress. Similarly, the role of the examiner's sex was determined influential in the interactional dynamics as well as in the attainment of the learner, albeit in variable cultural and person-level conditions. These variables, while researched in separate terms, are rarely taken as part of the combined effect within the context of a single empirical framework, especially within EFL speaking test scenarios. The lack of this makes current comprehension incomplete with reference to how speaking performance is determined under the mutual duress of affective, cognitive, and interpersonal variables. Bridging this lacuna, the present study integrates the above aspects in order to comprehend EFL speaking performance in finer terms with reference to application in terms of test architecture, examiner training, and selection of speaking topics in speaking proficiency tests.

CHAPTER III

METHODOLOGY

In order to examine anxiety level influence, gender of examiner, and speaking topics on EFL speaking performance, this chapter discusses research design, participants, data collection instruments, and procedures. The study adopts a mixed-method approach, which is an approach that integrates qualitative and quantitative approaches within a study. Quantitative data are essential for measuring, validating, and decision-making in product-led businesses. It can provide answers to questions such as user preference and preferred product features. Quantitative data may identify new questions that qualitative inquiries cannot provide answers to. Qualitative observations enable understanding of users who are pleased with a product or software for its feature set and users who churn or leave software (Kyne, 2021). The study adopts different data-gathering methods to examine learners' performance and perceptions regarding such factors as speaking examinations, interviews, and FLCAS surveys. The sampling process and data analysis approaches utilized to provide an explanation of findings are all explained in this chapter.

Research design

This study adopted a quasi-experimental mixed-methods design to investigate the effects of examiner gender, speaking topic type, and speaking anxiety on EFL learners' speaking performance. The design combined a quantitative strand, in which speaking performance and anxiety levels were measured under systematically varied conditions, with a qualitative strand that explored learners' perceptions through semi-structured interviews.

No treatment or instructional intervention was administered. Instead, the quantitative phase consisted of a single-session assessment in which all participants completed two speaking tasks—one on a gender-neutral topic and one on a gender-biased topic—under controlled conditions. The gender of the examiner (male or female) varied between participants through random assignment. Speaking anxiety was measured twice: once before the speaking tasks (pre-test anxiety) to capture baseline levels, and once after completion of both tasks (post-test anxiety) to assess any changes associated with the speaking experience.

The independent variables were:

1. Speaking anxiety level – measured pre-test and post-test.

2. Examiner gender – male or female.
3. Speaking topic type – gender-neutral vs. gender-biased.

The dependent variable was speaking performance, assessed using an analytical scoring rubric covering four criteria: fluency, accuracy, lexical resource, and coherence.

The qualitative phase directly followed the quantitative stage and was conducted with 16 students (8 males and 8 females) in the same session. Semi-structured interviews were used to gather participants' perspectives on speaking anxiety, the role of examiner gender, and how topic type influenced their comfort and performance.

This design allowed for controlled comparison of performance and anxiety levels across different conditions, while the qualitative data provided more profound insight into the patterns observed in the quantitative results.

Participants

A total of 46 undergraduate EFL learners initially took part in the study. However, due to incomplete data or exclusion criteria, the final sample was reduced to 46 participants who completed all stages of data collection, including the two speaking tasks and both anxiety assessments. Of these, 32 were female and 14 were male, reflecting the natural demographic distribution of the university's English language program. Although there was a noticeable gender imbalance, both male and female students were represented across all experimental conditions.All participants were enrolled in the university's English as a Foreign Language program and were considered to be at an intermediate proficiency level, approximately corresponding to B1–B2 on the Common European Framework of Reference for Languages (CEFR). This classification was based on internal placement procedures and instructor evaluations, even though standardized test scores were not available.

The intermediate level was considered suitable, as it ensured the speaking tasks were challenging but still within the learners' communicative abilities. Each learner completed two speaking tasks: one on a gender-biased topic and the other on a gender-neutral topic. To account for examiner-related factors, participants were randomly assigned to be evaluated by either a male or female examiner, which allowed examiner gender to be systematically controlled. In addition, to capture the learners' emotional responses, all participants

completed a foreign language speaking anxiety scale both before and after the speaking tasks.

Instruments

Four instruments were used to collect data for this study:

Speaking Tasks: Two speaking tasks were developed that would represent two types of topics—one gender-neutral and one gender-biased. The tasks were in similar structure and difficulty but were devised so as to yield spontaneous verbal responses. This would enable differences in ability to be linked with topic type and not with the difficulty of the tasks.

Anxiety Scale: A standardised Foreign Language Speaking Anxiety Scale created by Horwitz, Horwitz, and Cope (1986) was used with the purpose of measuring the level of anxiety of learners just prior to and after the speaking tasks. The scale included items that dealt with physical symptoms (e.g., nervousness, tension), emotional states (e.g., fear, discomfort), and cognitive worries (e.g., self-doubt, worry) related to speaking in a foreign language.

Performance Rubric: Speaking was scored with an analytic scoring rubric that comprised four criteria: fluency, accuracy, lexical resource, and coherence. The rubric allowed the inspection of numerous features of spoken language ability in depth.

Semi-Structured Interview: After the pilot phase, semi-structured interviews were carried out with 16 EFL students (8 males and 8 females). The interviews investigated the participants' experience with speaking anxiety, the way they see the effect of the sex of the examiner, and how they think various speaking tasks have impacted their ease and performance. The interviews gave rich qualitative information in support of the quantitative results.

Ethical Considerations

The respondents were informed of the purpose of the study, and they were permitted to participate voluntarily. The respondents remained anonymous and confidential, and ethical approval was obtained from the relevant institution.

Reliability and Validity of Instruments

So that the acquired data are valid and consistent, special care was given in this research concerning the reliability and the validity of the measuring instruments.

The scale of speaking anxiety used in this research is the Foreign Language Classroom Anxiety Scale (FLCAS), originally formulated by Horwitz, Horwitz, and Cope (1986). It has been widely employed in subsequent studies investigating speaking anxiety, such as Aida (1994), MacIntyre and Gardner (1991), Woodrow (2006), and Price (1991). Findings from these studies confirm that the FLCAS is a stable and reliable instrument for measuring the emotional states of foreign language learners.

The scale is implemented in order to evaluate cognitive, affective, and somatic features of language anxiety. The scale was identified in previous findings as being highly reliable with Cronbach's alpha values reported always over .80. The internal reliability in this study was re-calculated from the responses of the respondents, the resultant reliability coefficient being $\alpha = .86$ which is indicative of high internal consistency. This confirms that this instrument adequately measured this sample's target construct speaking anxiety.

To increase the speaking performance score reliability further, written responses from each test-taker were marked by a panel of four trained markers, two males and two females, for every administration of the test. Each of the four markers was precalibrated with benchmark specimens as well as the analytic rubric in order to assess fluency, accuracy, lexical resource, and coherence. Inter-rater reliability was established through the pilot score-setting session with the following group discussion in order to settle the differences and standardize the scoring practices. The ultimate score on every task was calculated as the average of the scores from four raters in order to reduce the biases of the individuals as well as standardize the scoring practices better.

Additionally, using a gender-balanced panel of raters also supported face validity and fairness of the procedure of assessment. This was especially required because one of this research's independent variables was the gender of the examiners. Standardization of procedure and deliberate inclusion of both male and female assessors ensured that the procedure of assessment reflected design priorities of this research and maximized its findings' acceptability.

Data Collection Procedure

Data collection took place within one session per participant, with a standardised order to ensure consistency. Participants were greeted on arrival and informed of the purpose of the study with assurances about the confidentiality of their answers prior to giving informed consent. The session commenced with the administration of a pre-test anxiety questionnaire to set up baseline speaking anxiety levels. Participants next did two speaking tests, a gender-neutral and a gender-biased one respectively, presented in counterbalanced order to reduce order effects. The gender of the examiner was varied among participants by random assignment to either male or female examiner, with both tasks being undertaken by the same examiner per participant. Participants moved straight on to the second task following completion of this first one, and both performances were audio-recorded for later analysis using an analytical rubric to assess fluency, accuracy, lexical resource, and coherence. Participants then completed at this point the same anxiety questionnaire as that given at the commencement of the session to identify any changes to anxiety level. Finally, 16 participants (8 males and 8 females) took part in a semi-structured interview, immediately following on from the post-test measure, during which they commented on their experiences during the speaking tests. The interviews probed their feelings about anxiety, comfort levels with varying examiner gender, and difficulty with topics they received, as well as what they did to cope with anxiety. The interviews were all audio-recorded with transcription verbatim for later thematic analysis.

Data Analysis Procedures

Data analysis for this study was conducted using the Statistical Package for the Social Sciences (SPSS). Descriptive and inferential statistics were applied to respond to research questions as well as test hypotheses on speaking topic type, gender of examiner, and anxiety on EFL speakers' performance.

Descriptive statistics with means, standard deviations, and minimums and maximums were calculated for gender-neutral and gender-biased speaking task performance by students and pre- and post-test anxiety levels. They provided a summary of performance trends and distribution of anxiety at the sample level.

To determine whether performance was significantly different for speaking topic types, and to explore shifts in anxiety levels, paired-samples t-tests were conducted. The tests paired each student's performance on two related conditions (gender-biased vs. gender-neutral; pre-test vs. post-test anxiety), with individual variation controlled.

To identify anxiety and speaking performance correlation, Pearson correlation analysis was applied. It assessed strength and direction of association between pre- and post-tests of anxiety level of learners and their speaking performance regarding two types of topics.

We applied repeated measures analysis of variance (ANOVA) in order to uncover the effect of the type of speaking topic on the learners' performance, with the within-subject variation considered. The test formed Topic Type (GN vs. GB) as the within-subjects variable and tested whether the learners' performance significantly differed under different types of tasks.

To examine the effect of examiner sex with pre-test anxiety as a possible confounding variable statistically controlled, an Analysis of Covariance (ANCOVA) was applied. The dependent variable was the difference in scores on gender-neutral as compared with gender-based material; the independent variable was the sex of the examiner; and the covariate was pre-test anxiety. Through this procedure, we were able to examine examiner sex effects more accurately after statistically controlling the differences in affective states.

Prior to conducting ANCOVA, assumption tests were conducted in order to verify that data had the desired requirements in order for correct interpretation. These consisted of normality tests of residuals, homogeneity of variances (from Levine’s Test), linearity, and homogeneity of regression slopes (using interaction tests). All the assumptions were satisfied.

After the quantitative analysis was undertaken, the interviews were investigated through the use of thematic analysis in order to identify common patterns and underlying meaning the participant responses. Beginning with the careful reading of all of the interview notes, multiple times, in order to familiarize oneself with the data and gain insight into the initial ideas. Then, coding was undertaken seeking out portions of text, such as phrases or complete sentences, that indicated significant ideas, emotions, or experiences. After all data had been coded, the codes were reviewed for patterns and associations, and similar or related codes were collated together in order to create themes. This was done through checking that

each theme accurately represented the information that it was based on and that the themes differed enough from one another. Some themes were narrowed, combined, or divided as required in order to gain clarity and cohesiveness. Each theme was then succinctly defined and labeled in order to reflect its core concept. Finally, eight major themes were identified that are reported in concord with the interview questions that guided the interview, offering an ordered and significant meaning of the qualitative data.

CHAPTER IV

FINDINGS

This study employed a combination of descriptive and inferential statistical analyses to investigate the effects of examiner gender, speaking topic type, and test anxiety on EFL learners' speaking performance. Descriptive statistics were first used to summarize the learners' mean scores for gender-neutral and gender-biased topics, along with their pre- and post-test anxiety levels. These initial results provided a broad overview of performance patterns and emotional states during the assessment process.

To assess whether differences in speaking performance between topic types were statistically significant, paired-samples t-tests were conducted. This analysis revealed that learners achieved significantly higher scores on gender-neutral topics, although no significant change was observed in anxiety levels from pre- to post-test. Pearson correlation analysis was then applied to explore the relationship between anxiety and speaking performance. The findings indicated a significant negative correlation, with higher anxiety levels associated with lower performance, particularly on gender-biased topics, suggesting that anxiety played a meaningful role in shaping outcomes.

To investigate performance differences within subjects on different types of topics, a Repeated Measures ANOVA was later applied. Taking topic type (gender-neutral versus gender-biased) as a within-subjects factor, this analysis maintained a significant main effect with subjects performing significantly better on gender-neutral speaking tasks. This procedure increased internal validity by controlling subject differences on performance variability.

To assess the influence of examiner gender controlling for learners' anxiety, an Analysis of Covariance (ANCOVA) was used. ANCOVA was selected because it can be used to compare group means adjusting for a continuous covariate, which in this case was pre-test anxiety. The error variance associated with emotional differences was thereby diminished by this approach to make the findings more precise. All assumptions necessary for ANCOVA, such as normality of residuals, linearity, homogeneity of variances, and homogeneity of regression slopes were tested and found to hold.

Rationale for Using ANCOVA Instead of ANOVA

The choice of using ANCOVA, instead of an ordinary ANOVA, was informed by the necessity of controlling the confounding variable of learners' anxiety levels when analyzing the effect of examiner gender on speaking performance. As much as ANOVA is capable of picking up mean differences among groups, it does not take into account continuous variables that can affect the dependent variable. ANCOVA puts the strengths of ANOVA and regression together, allowing the incorporation of the use of a covariate in order to adjust the group means due to prior differences. Anxiety in this study was revealed through the negative correlation with the performance in the gender-biased tasks; consequently, the controlling of the anxiety was critical in order not to blame the difference in the performance based solely on the examiner's gender. By keeping the anxiety constant, the ANCOVA gave more precise estimation of the effect of examiner's gender, better statistical power, and enhanced interpretability.

Overall, the addition of descriptive statistics, t-tests, correlation analysis, repeated measures ANOVA, and ANCOVA permitted an in-depth and methodologically rigorous examination of the research questions, appropriately controlling for confounding variables and bolstering the validity of the results.

Table 5. Descriptive Statistics of Performance and Anxiety (SPSS Style)

Variable	N	Mean	Std. Deviation	Minimum	Maximum
Score on GB	46	3.83	0.94	1.80	5.00
Score on GN	46	4.12	0.79	2.20	5.00
Anxiety (Pre)	46	73.37	14.42	50.00	121.00
Anxiety (Post)	46	71.89	15.81	41.00	122.00

Descriptive statistics offer clear indications of speaking ability differences along with levels of anxiety among the test-takers. The students performed better on average on gender-neutral (GN) material (M = 4.12, SD = 0.79) compared to gender-biased (GB) material (M = 3.83, SD = 0.94), while the large standard deviation in GB material is indicative of more variation in speaking ability on GB tasks.

Regarding anxiety, the mean pre-test score was 73.37, while the mean post-test score decreased slightly to 71.89. Although this decline points to a modest reduction in anxiety following the speaking exam, the relatively large standard deviations (14.42 and 15.81,

respectively) highlight considerable individual differences in how learners experienced and responded to test-related stress.

These descriptive results provide an initial overview of performance patterns and emotional states, laying the groundwork for the subsequent inferential analyses assessing these trends' statistical significance.

Table 6. Paired Samples Test (SPSS Style)

Paired Differences	Mean	Std. Deviation	Std. Error Mean	t	df	Sig.(2tailed)
Score on GN – Score on GB	0.29	0.59	0.096	-3.00	45	.004
Anxiety (Post) – Anxiety (Pre)	-1.48	16.50	2.45	0.90	45	.371

Two paired-samples t-tests were conducted to examine differences in learners' speaking performance across topic types and changes in their anxiety levels.

For speaking performance, a statistically significant difference emerged between scores on gender-neutral and gender-biased topics, $t(45) = -3.00$, $p = .004$. Learners performed notably better on gender-neutral topics (M = 4.12) than on gender-biased topics (M = 3.83), with a mean difference of 0.29. This finding indicates that topic type had a measurable and meaningful effect on speaking performance.

In contrast, the analysis of anxiety levels revealed no statistically significant difference between pre-test and post-test scores, $t(45) = 0.90$, $p = .371$. Although mean post-test anxiety (M = 71.89) was slightly lower than mean pre-test anxiety (M = 73.37), the change was minimal and statistically unreliable. This suggests that participation in the speaking test did not significantly shift learners' anxiety levels.

Pearson Correlation Coefficients (N = 46)

Table 7. Correlations Between Anxiety and Speaking Performance (SPSS Style)

Variables	Score GB	Score GN
Pre-Anxiety	-0.33*	-0.24
	(.024)	(.110)
Post-Anxiety	-0.30*	-0.14
	(.045)	(.368)

Note: Values in parentheses represent **Sig. (2-tailed)**. *Correlation is significant at the 0.05 level (2-tailed).

Pearson correlation analysis was conducted to examine the relationship between learners' anxiety levels (pre- and post-test) and their speaking performance on gender-biased (GB) and gender-neutral (GN) topics. The results revealed a significant negative association between anxiety and performance on GB tasks. Specifically, pre-test anxiety correlated negatively with GB scores (r = -0.33, p = .024), and post-test anxiety also showed a significant negative correlation (r = -0.30, p = .045). These findings indicate that higher anxiety levels, both before and after the exam, were linked to lower performance on gender-biased speaking tasks.

In contrast, the correlations between anxiety and GN performance were weaker and not statistically significant. The relationship between pre-test anxiety and GN scores was negative but non-significant (r = -0.24, p = .110), and the correlation between post-test anxiety and GN scores was similarly weak and non-significant (r = -0.14, p = .368). This pattern suggests that while anxiety had a consistent detrimental effect on performance in gender-biased topics, its impact on gender-neutral topics was less pronounced and not statistically reliable.

Table 8. Tests of Normality – Shapiro-Wilk

Variable	Statistic	df	Sig.
ANCOVA Residuals	0.976	46	.442

The Shapiro–Wilk test for normality was non-significant (p = .442), indicating that the residuals of the ANCOVA model followed a normal distribution. This result confirms that the assumption of normality was satisfied.

Table 9. Levene's Test of Equality of Error Variances

Dependent Variable	F	df1	df2
Score_Diff (GN–GB)	0.041	1	44

Levene's test for equality of variances was non-significant (p = .840), indicating that variances were equal across examiner gender groups. This result confirms that the assumption of homogeneity of variance was met.

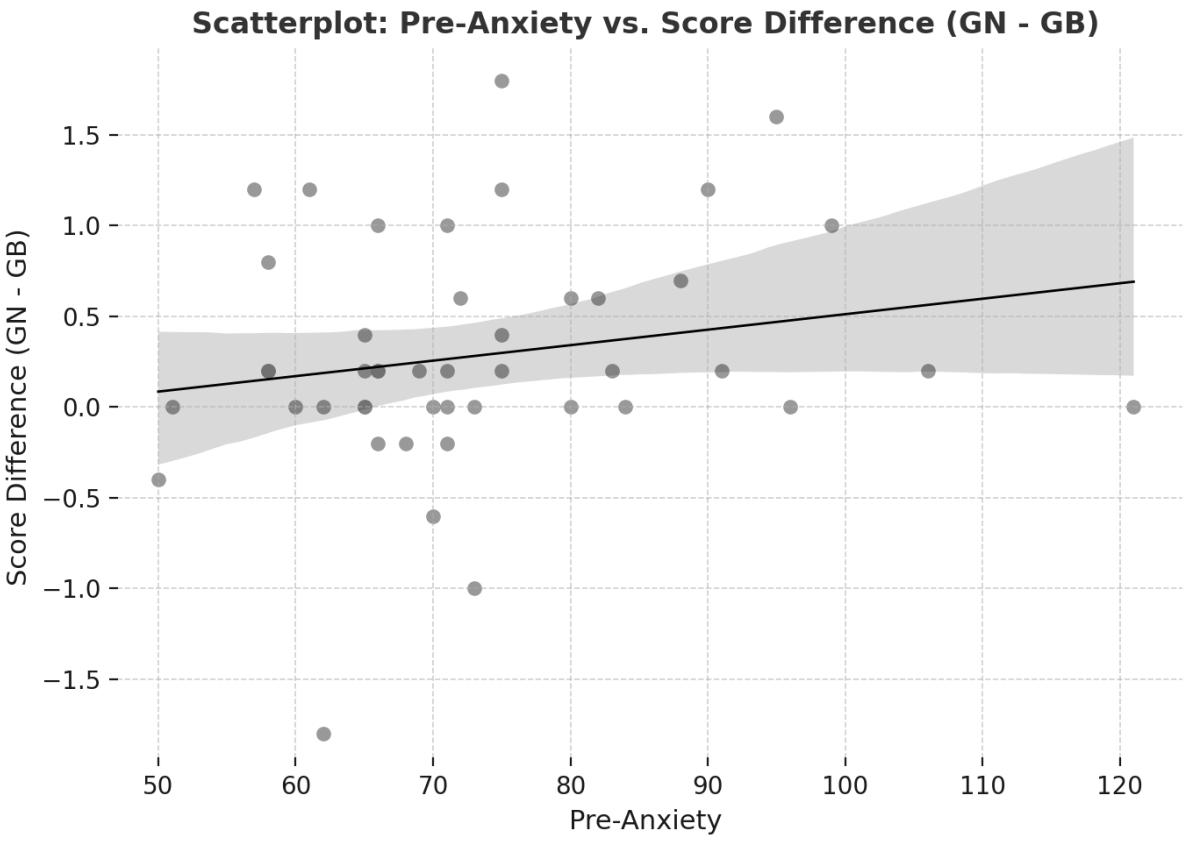

Figure 1. Scatterplot: Pre-Anxiety Vs.Score Difference (GN - GB)

Linearity Assumption

Scatterplots were generated to visually examine the relationship between covariate (pre-test anxiety) and the outcome variable (difference score on speaking tasks between gender-biased and gender-neutral speaking tasks) to check for the linearity assumption of ANCOVA. Figure 2 illustrates that a clear downward-sloping linear trend was found for the pre-anxiety versus score difference scatterplot, which shows that pre-anxiety levels were higher as differences on speaking scores were smaller. The data points tended to follow a downward-sloping regression line as opposed to a curvilinear or random shape.

By contrast, the post-anxiety versus score difference scatterplot displayed a weaker and more scattered distribution. Although a slight negative trend was visible, it lacked the consistency in the pre-anxiety plot.

The linearity assumption was found to be met because only pre-anxiety was used as a covariate under the ANCOVA model, and its relationship with the outcome variable was linear as seen visually. The visual evidence accompanied by pre-anxiety theoretical justification as a beginning point indicator of affective state justifies its inclusion as a covariate. Hence, the statistical control applied using ANCOVA was methodologically sound with data structure justification.

Table 10. Test of Homogeneity of Regression Slopes

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Examiner_Gender	0.000	1	0.000	0.103	.750
Pre-Anxiety	4650.154	1	4650.154	2.201E30	.000
Examiner_Gender × Pre-Anxiety	0.000	1	0.000	0.212	.648
Error	0.000	45	0.000		

The covariate-examiner interaction (pre-test anxiety) and independent variable (examiner gender) was tested to determine whether or not association between anxiety and difference in speaking performance (Score_Diff) varied by groups of examiners. Interaction term (Examiner Gender × Pre-Anxiety) was found to be non-significant ($p > .05$), which implies that anxiety effect on score difference was equal for male and female examiners. The outcome meets the assumption of homogeneity of regression slopes, which verifies that relationship between covariate and dependent variable does not vary by group. Therefore,

proceeding with ANCOVA was okay because this crucial assumption for model validity was achieved.

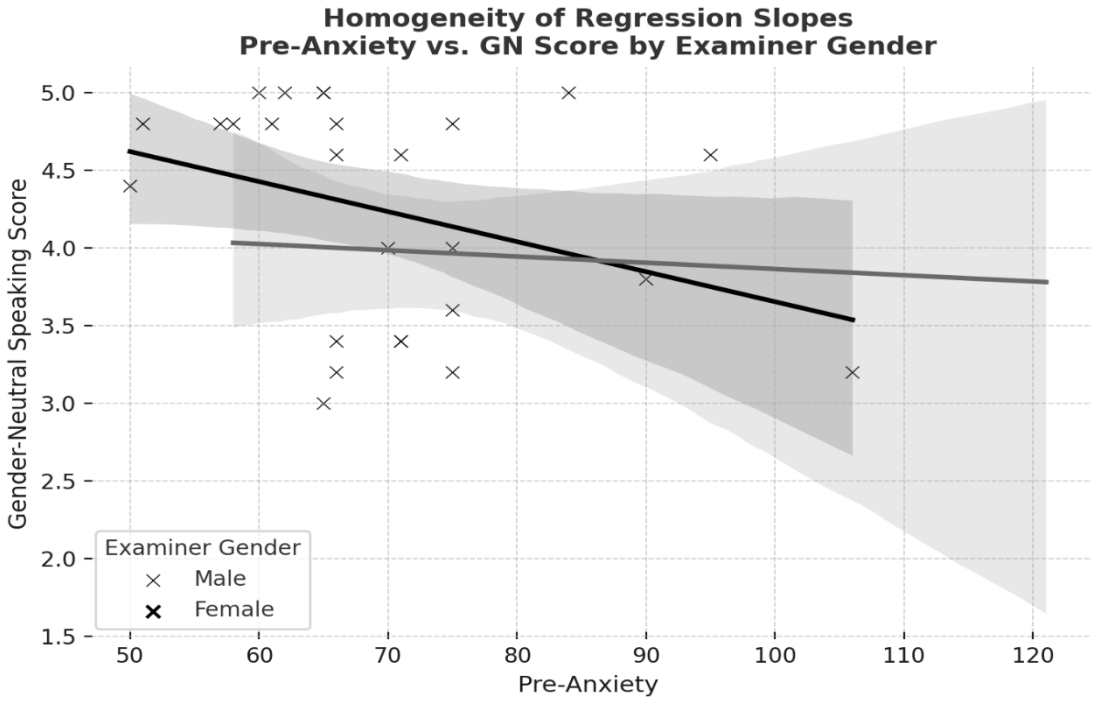

Figure 2. Homogeneity of Regression Slopes Pre-Anxiety Vs. GN Score by Examiner Gender

To check on the assumption of homogeneity of regression slopes for ANCOVA, a scatterplot was generated to examine the association between pre-test anxiety (covariate) and the difference in score (GN − GB) with examiner gender as a separator. As Figure X reveals, the male and female examiner regression lines are on roughly the same track and are roughly parallel to each other, which indicates that the influence of anxiety on performance difference holds steady for groups. The visual evidence is consistent with the outcome of the statistical test for interaction, wherein the term Examiner Gender × Pre-Anxiety was not significant ($p > .05$). Together, these findings verify that the assumption of homogeneity of regression slopes holds true and thereby warrants conducting ANCOVA to study the influence of examiner gender on anxiety as a control.

Table 11. Tests of Between-Subjects Effects (ANCOVA Summary)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Examiner Gender	2.193	1	2.193	5.97	.019
Pre-Anxiety	1.462	1	1.462	3.98	.052
Error	15.805	43	0.368		

An ANCOVA was conducted to examine the effect of examiner gender on the difference in speaking scores between gender-neutral and gender-biased topics, with pre-test anxiety included as a covariate. The results indicated a statistically significant main effect of examiner gender on score difference, $F(1, 43) = 5.97$, $p = .019$, suggesting that examiner gender influenced learners' relative performance across topic types even after controlling for anxiety levels.

The covariate pre-test anxiety had a marginally significant effect on speaking performance, $F(1, 43) = 3.98$, $p = .052$, suggesting a trend in which higher levels of anxiety were related to smaller differences in speaking performance. However, this did not pass conventional levels of statistical significance.

Adjusted means scores reported that those examined by male examiners had a higher adjusted diff score ($M\_diff = 0.44$) than those tested by female examiners ($M\_diff = 0.09$). It can be inferred that those tested by a male examiner tended to score relatively higher on gender-neutral issues, suggesting that there might be an interaction between topic types and examiner properties that influences speaking performance.

Repeated Measures ANOVA

We applied a Repeated Measures ANOVA here in order to discover whether speaking performance differed significantly from one type of speaking task to another, that is, gender-biased (GB) and gender-neutral (GN) topics. As both participants attempted both kinds of tasks, Topic_Type was viewed as a within-subjects variable, enabling comparison at the level of each person's performance with that under the other condition. The procedure is well

tailored to within-participant design as the variation at the individual level is controlled and the effect of type of task per se is dissociated.

Analysis indicated a significant main effect of Topic_Type, with performance shifting in the usual way across conditions with increasing scores under the condition of gender-neutral topics. The finding suggests that the speaking task structure or speaking-task nature may directly affect the test performance through its effect on comfort level, topic familiarity, or cognitive processing demand at the test.

Table 12. Repeated Measures ANOVA (Topic Effect Only)

Source	df	F	Sig.
Topic_Type	1	9.00	.004
Error	45	—	—

The repeated measures ANOVA indicated a significant main effect of speaking topic type on performance, $F(1, 45) = 9.00$, $p = .004$, thus verifying that learners significantly outperformed other learners in speaking scores on gender-neutral topics relative to gender-biased topics. The within-subjects design maximized statistical power through statistical control of individual variation in language proficiency, confidence, as well as other personal qualities, in order to more sensitively contrast tasks types.

Comparison of Repeated Measures ANOVA and ANCOVA

In order to examine the performance in topic types, at the beginning a repeated measures ANOVA was applied in order to compare scores in gender-neutral (GN) and gender-biased (GB) topics in the same individuals. The results showed that there was a significant main effect in topic type, indicating that the learners significantly outperformed in GN tasks. The finding emphasized the influence of the nature of the contents of tasks in speaking ability. However, while repeated measures ANOVA cancels out within-subject variability, continuous variables such as anxiety that affect the performance continue to remain uncontrolled.

To overcome this weakness, an ANCOVA was then conducted using pre-test anxiety as the covariate. The application of this procedure permitted more precise examination of the magnitude of the effect of sex of examiner on speaking performance differences while statistically controlling the possible confounding variable of anxiety. The findings of the

ANCOVA were that the effect of sex of examiner on differences in performance continued significantly when controlled at the level of anxiety, but that anxiety per se had only marginal effects. By reducing the error variance and statistically controlling extranea variables, the ANCOVA enhanced the explanatory power of the analysis and permitted more precise estimation of the true effect of sex of examiner.

The two analyses conducted offered each other mutually supplementary findings: the repeated measures ANOVA reproduced the significant topic type effect on performance. At the same time, the application of ANCOVA revealed the examiner sex effect as well as the partial effect of anxiety. The joint effort of the two analytical endeavors strengthened the result's validity as well as added depth to the understanding of what enters into EFL learners' speaking proficiency.

Summary of Findings

The present chapter presented the results of statistical analysis exploring speaking topic type, sex of the examiner, and test anxiousy as predictors of speaking proficiency of EFL learners. Descriptive and inferential statistical procedures, descriptive statistics, paired-samples t-tests, Pearson correlation procedures, repeated measures ANOVA, and ANCOVA were employed in responding to the research questions with clearness and completeness.

Descriptive statistics indicated that the students performed better in the gender-neutral (GN) areas compared to the gender-biased (GB) areas. The post-test level of anxiety was slightly lower than the pre-test level, but this was little and was statistically insignificant. Paired-samples t-tests also reiterated that there was significant outperformance in GN areas ($p = .004$). However, the comparison of the level of anxiety in the post- and pre-test was insignificant ($p = .371$), so that the speaking tasks per se did not significantly reduce nor significantly increase test anxiety.

To investigate the relationship between anxiety and performance, Pearson correlation analyses were conducted. The findings revealed that higher levels of anxiety were significantly associated with lower performance on GB tasks, with both pre- and post-test anxiety negatively correlated with GB scores ($r = -0.33$, $p = .024$; $r = -0.30$, $p = .045$). In contrast, no significant correlations were observed between anxiety levels and GN scores.

This pattern suggests that anxiety was more detrimental when learners engaged with topics that were potentially sensitive or less familiar.

A repeated measures ANOVA, with *Topic_Type* (GB vs. GN) as a within-subjects factor, confirmed a significant main effect of topic type, $F(1, 45) = 8.99$, $p = .004$, indicating that learners consistently performed better on GN tasks. The within-subjects design controlled for individual variability, strengthening the conclusion that task content alone influenced performance.

Before conducting the ANCOVA, all statistical assumptions—normality of residuals, homogeneity of variances, linearity, and homogeneity of regression slopes—were tested and met. The ANCOVA assessed whether examiner gender influenced performance differences between topic types after controlling for pre-test anxiety. The results showed a significant effect of examiner gender, $F(1, 43) = 5.97$, $p = .019$, with learners examined by male examiners exhibiting a greater performance difference in favor of GN topics compared to those examined by female examiners. Although pre-test anxiety approached statistical significance ($p = .052$), it did not independently predict performance differences at the conventional threshold.

In summary

The findings indicate that learners performed significantly better on gender-neutral topics than on gender-biased topics, a result confirmed by the repeated measures ANOVA, which controlled for within-subject variability. Examiner gender also significantly impacted the difference in speaking performance between topic types, with performance patterns varying according to the examiner's gender. Furthermore, higher anxiety levels were associated with lower performance, particularly on gender-biased tasks, highlighting the detrimental role of anxiety in specific assessment contexts. All assumptions for ANCOVA were satisfied, ensuring the validity and robustness of these results.

Qualitative Findings

This paper reports qualitative findings from semi-structured interviews with 16 EFL learners (eight males and eight females). Learners' experiences of speaking anxiety, examiners' gender as perceived influence on performance, and topic type as influence on performance were investigated through interviews. The data were theme-coded according to

Braun and Clarke's (2006) six-phase thematic analysis procedure. Initial codes were condensed and related to six overarching themes with supporting illustrative quotations. Where necessary, figures and tables are provided to present key patterns of respondents' answers visually.

Theme 1: Pre-speaking and During-speaking Anxiety

Most participants reported heightened anxiety before speaking tasks, often linked to insufficient preparation, fear of unexpected questions, or low confidence in language abilities. This anxiety was expressed physically (stress, confusion, nervousness) and cognitively (racing thoughts, fear of judgment).

"Before a speaking task, I tend to be more anxious than normal... maybe something unexpected can happen, so it is kinda scary." (Male 1)

"Before speaking, I generally feel a little bit confused and stressed." (Male 5)

"Before speaking, I feel anxious and excited, I start thinking about what I'm going to say in order not to let the others laugh at me." (Male 8)

Some participants became more relaxed during the task, while others—particularly those with higher social anxiety—remained anxious.

"While speaking, I tend to be so relaxed usually because I've already started talking and doing good." (Male 1)

"During the speaking, I feel afraid that someone would make fun of me." (Male 8)

"During the speaking task I feel more anxious." (Female 4)

Table 1 summarises the frequency and percentage of participants reporting pre-speaking anxiety, changes in anxiety during speaking, and use of breathing/relaxation techniques.

Table 13. Frequency and percentage of key anxiety-related responses

Theme	Frequency	Percentage (%)
Pre-speaking Anxiety Reported	16	100.0
During-speaking Anxiety Decreases	6	37.5
During-speaking Anxiety Increases or Persists	6	37.5
Uses Breathing/Relaxation Techniques	5	31.2

These results indicate that while pre-speaking anxiety is nearly universal among participants, their experiences during the task diverge, with some adapting to the situation and others struggling to regulate their emotional state.

Theme 2: Coping Strategies for Speaking Anxiety

Participants described various strategies for managing speaking anxiety, grouped into four main categories: Distraction Techniques, Breathing and Relaxation, Focus on the Task, and Self-Encouragement.

Distraction Techniques included avoiding eye contact, fidgeting, or imagining oneself in a different environment.

"I try to imagine myself in a different place, I avoid eye contact with the audience." (Male 3)

Breathing and Relaxation involved deep breathing, meditation, or stepping outside for fresh air.

“I do meditation and I try to relax.” (Male 6)

Focus on the Task meant concentrating on the topic rather than the audience or personal feelings.

"If I prepare before I speak, it makes me more comfortable; if I know the subject before, it will be good." (Female 2)

Self-encouragement involved positive self-talk and affirmations.

"While I'm speaking, I take a deep breath and I say to myself that everything is gonna be okay, you can do it." (Male 5)

Some participants also acknowledged that overthinking intensified their anxiety.

"I think too much before speaking." (Female 1)

Figure illustrates the distribution of coping strategies, showing breathing/relaxation as the most common approach.

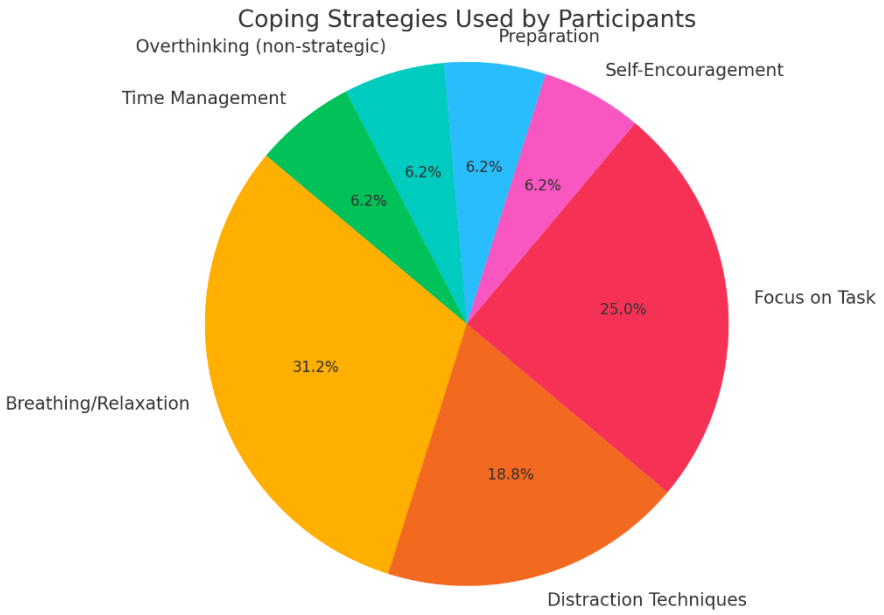

Figure 3.Coping strategies used by participants

Theme 3: Perceived Impact of Anxiety on Speaking Performance

Most participants reported that anxiety impaired their speaking performance by causing word loss, grammatical mistakes, reduced fluency, or decreased confidence.

"It hinders my potential and my capacity to speak better." (Male 3)

“I create wrong sentences, my voice always shakes.” (Female 1)

"It affects me so bad that I can't express myself correctly." (Male 6)

A few learners noted that anxiety had little lasting effect once they began speaking.

"Anxiety doesn't affect my overall speaking performance… when I start speaking, it goes totally off the track." (Male 1)

These findings suggest that while some learners can self-regulate, for most, anxiety is a significant barrier to optimal performance.

Theme 4: Situational Triggers of Anxiety

Participants identified several situations that heightened anxiety, including fear of negative evaluation, lack of preparation, formal exam settings, technical issues, and disengaged listeners.

“Seeing others fail made me anxious.” (Male 1)

“I forgot everything and started thinking about why they laughed at me.” (Male 8)

“There was a presentation… I couldn’t find the PowerPoint file… I had to speak spontaneously. I really felt so nervous and anxious that day.” (Male 7)

These accounts show how external conditions and social reactions can amplify anxiety, often undermining performance.

Theme 5: Examiner’s Gender and Comfort Level

Most male participants said examiner gender did not influence their comfort, while some preferred female examiners for their perceived empathy and reduced criticism.

“I feel more comfortable when I’m speaking with a female examiner because they understand me and listen to me.” (Male 8)

Female participants were more divided—some preferred female examiners for their perceived understanding, while others found male examiners more challenging.

"Speaking to a male is more challenging." (Female 2)

"Speaking to a female examiner is easier for me." (Female 4)

Figure 2 summarises these preferences by participant gender, illustrating that while neutrality was common, notable subsets expressed gender-based preferences.

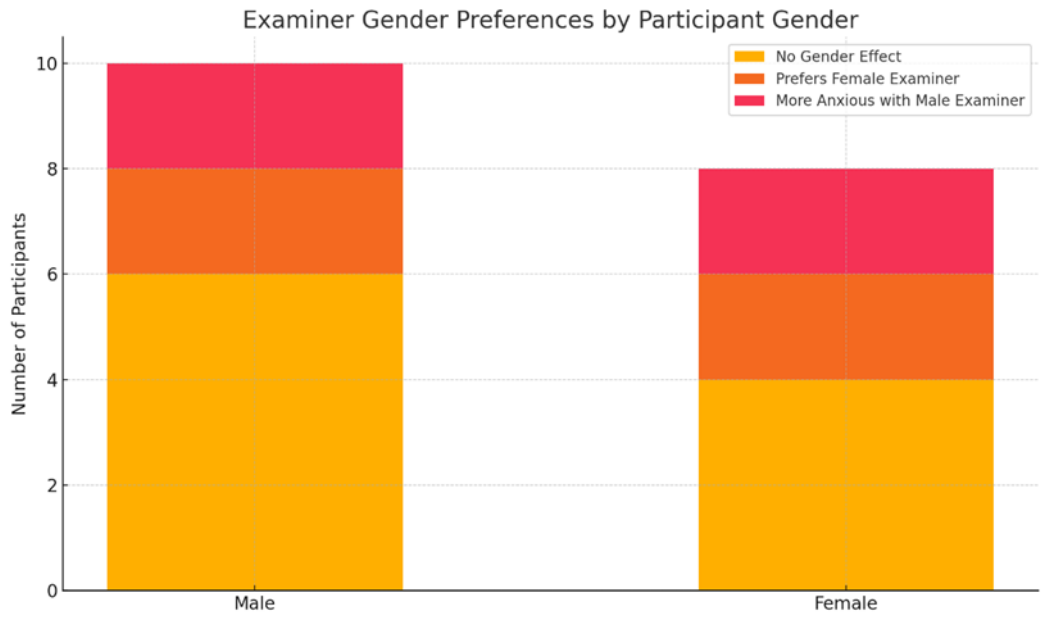

Figure 4. Examiner's gender preferences by participant gender

Theme 6: Topic Type and Anxiety – Gender-Biased vs. Gender-Neutral

Gender-neutral topics (e.g., technology, hobbies) were generally perceived as easier and less stressful.

"I feel more confident, more talkative, and more relaxed." (Male 2)

"I feel comfortable and more relaxed." (Female 1)

Gender-biased topics (e.g., gender roles, feminism) often increased anxiety, especially when participants feared offending others or lacked knowledge.

"*I feel nervous because I don't have ideas about these topics."* (Female 5)

“I feel anxious because I’m not educated about these topics.” (Male 2)

Some participants linked topic type and examiner gender, noting discomfort discussing gender-biased topics with certain examiners.

"I feel anxious when discussing gender-biased topics with the male examiner, but I feel relaxed… with the female." (Female 5)

These findings indicate that both topic type and examiner gender can interact to shape comfort levels and, ultimately, speaking performance.

Summary of Qualitative Findings

Qualitative data reveal that EFL speaking anxiety is determined by a complex set of personal, situation, and task factors occurring together. While coping strategies are effective for certain learners to minimize anxiety, they are ineffective for others, particularly under socially sensitive or high-stakes situations. Gender of examiner and topic type are found to be substantial contextual factors which come together to influence comfort and performance. These qualitative observations complement the quantitative analyses to provide an insightful perspective on EFL speaking outcome-determining factors.

CHAPTER V

DISCUSSION

The following Chapter interprets and critically evaluates findings of this research against the framework of its four research questions with special interest in contrasting findings with findings of prior works on applied linguistics, second language acquisition (SLA), and language tests. The core findings are discussed with the related empirical works and theoretical positions with the points of convergence, contrast, and potential explanations of contrast.

The study found that the speakers fared better when speaking on the gender-neutral topic compared with the topic that was biased. A repeated-measures ANOVA confirmed this, indicating that there was a significant difference in speaking ability in terms of the topic type. Confirmatory evidence is the idea that what the task is about extensively dictates what the learner will produce, an argument long established (Kim, 2009; Iwashita et al., 2008). Further evidence from more recent dates is in agreement. For instance, Jankowiak, Błaszczak, and Dziubalska-Kołaczyk (2024) confirmed that culturally sensitive or socially related topics put the learner on edge and reduce fluency. However, Mora, Treffers-Daller, and Alario (2025) believe that if the topic is well-known, the learning is easier and there is more use of vocabulary.

In terms of anxiety, the results showed only a small, not meaningful, decrease from before to after the test. This aligns with what earlier research claimed about language anxiety being relatively stable over time (Horwitz et al., 1986; Woodrow, 2006). But this is different from Cheng et al. (1999), who observed significant improvements after a long period of preparation. That shows that short-term exposure might not be enough. Recent studies say that keeping up with practice and using smart ways to deal with anxiety is important. For example, Szyszka (2024) found that learners on their own use techniques like relaxation and talking to themselves. However, Garcia, Lopez, and Chan (2024) found that these methods might not work well when the task has a lot of personal or social meaning. The study here shows that even though anxiety didn’t drop much, learners are becoming more aware of how to handle their anxiety through thinking about strategies.

The correlation analysis showed that anxiety had a strong negative effect on performance when the task involved gender-biased topics, but this effect was less strong for

gender-neutral topics. This result matches Krashen's (1982) Affective Filter Hypothesis and MacIntyre and Gardner's (1994) research on how anxiety can interfere with thinking. New studies also back up the idea that what a task is about can influence how anxiety affects performance (Mora, Treffers-Daller, & Alario, 2025). Although task difficulty is often discussed in studies, these findings show that social sensitivity can also greatly affect performance. This means that those who design tests should take both emotional and mental demands into account when creating tasks.

The ANCOVA results showed that the gender of the examiner had a significant effect on how well learners performed in speaking, especially when looking at how they interacted with different types of topics. Learners who were evaluated by male examiners showed more differences in their performance between gender-neutral topics and those that had a gender bias, compared to learners evaluated by female examiners. This matches what other studies have found about how examiners can influence results (O'Loughlin, 2002; Lumley & Sullivan, 2005) and shows the importance of the social and emotional parts of speaking (Tantucci, 2025). The findings of this study show that it's important for examiners to be trained and aware so that personal relationship factors don't create unfair bias.

Finally, the coping methods that learners used, such as talking to themselves positively, relaxing, and practicing, are similar to the strategies that Oxford (1990) described for managing emotions. Szyszka(2024) also supports this with findings about how learners control their emotions. Even though these strategies were used, learners still felt anxious, which suggests they might not be very effective in high-pressure situations right away. New technology, like virtual reality, has shown promise in helping reduce speaking anxiety by offering a safe and low-pressure way to practice (Thorpe, 2025). Although these tools weren't used in this study, they suggest that combining traditional methods with technology-based practice could be a good area for future research and teaching.

CHAPTER VI

CONCLUSION AND IMPLICATIONS

This study examined the influence of three related variables: anxiety of learners, gender of tester, and gender-neutral versus gender-biased type of speaking theme, on speaking proficiency of English as a Foreign Language (EFL) learners. Speaking proficiency during test conditions is regulated by something more than proficiency at a linguistic level; situation-related as well as psychological factors can aid or obstruct learners to speak eloquently with confidence (Tsegaye & Kassa, 2023; Zhang & Wang, 2022).

The primary aim was to determine whether topic type influences speaking performance, focusing on whether gender-biased topics affect learners differently from gender-neutral ones. A second aim was to explore whether examiner gender impacts performance, especially on sensitive or personal topics. The study also sought to identify changes in learners' anxiety levels before and after a speaking task, and to examine the relationship between anxiety and speaking scores.

Recognition of such factors is crucial in modern language testing, wherein test anxiety is reported as a significant barrier to the best possible performance, particularly in speaking tasks (Ali & Rezaei, 2022; Karimi & Zhaleh, 2024). Test results may be compromised as a function of topic familiarity and test examiner characteristics (e.g., local culture familiarity), confounding fairness and validity (Chen & Lee, 2023). Examination of these aspects sheds insight upon this study's attempts at fairly constructed tests, while practical applications of this work extend toward reducing anxiety and improving speaking in EFL contexts

Results showed some interesting patterns. First, the students scored significantly better on gender-neutral compared to on gender-biased topics, indicating that neutral topics minimize affective load and enable people to perform with more confidence and with more fluency. Second, levels of anxiety moderated slightly but significantly from pre-test through post-test, indicating that short exposure to speaking tasks will decrease anxiety but may not decrease anxiety in any significant manner. Third, levels of anxiety negatively correlated strongly with performance on gender-biased topics, indicating that high levels of anxiety may interfere with performance when performance entails processing sensitive information.

Fourth, the gender of the examiner significantly moderated the performance on both gender-biased as well as on gender-neutral topics, although in a lesser extent, indicating that what people think the examiner as part of society will be like will affect the way people will perform under high-stakes orals. Lastly, qualitative findings reported that the students strove more at the use of preparation, cognitive re-framing of topics, as well as at the use of relaxation strategies so as to minimize anxiety.

These results uncover subtle affective, social, and lexical dynamics in speaking tests. The findings on topic type and sex of examiner mirror that speaking test fairness should under no circumstances be assumed based on test design and examiner training. From the pedagogical point of view, they stress the significance of topic selection as inclusive and culture-sensitive topics as well-known as possible among all the learners (Chen & Lee, 2023). They stress the significance of offering anxiety-management strategies to the learners and designing non-intimidating speaking settings with minimized pressure (Ali & Rezaei, 2022).Theoretically, this study contributes to work on the affective filter hypothesis (Krashen, 1985) and sociocultural perspectives on language learning, which recognise that performance is shaped by the interplay between individual emotions, social relationships, and task conditions (Karimi & Zhaleh, 2024). The results challenge the assumption that oral assessments are entirely objective, suggesting that examiner characteristics and topic choice may unintentionally influence outcomes. This supports calls for more robust fairness validation in speaking test design and for assessment models that integrate affective and social dimensions alongside linguistic competence (Zhang & Wang, 2022).

Practically speaking, the study presents sound recommendations for teacher preparation and examiner training. Examiners should be trained to follow an encouraging neutral style of interaction to reduce anxiety and performance inhibitions. Institutions may also wish to vary examiner gender representation or give students limited freedom of choice over topics to minimize possible bias. Furthermore, as learners indicated they used a variety of strategies to cope with anxiety, the addition of explicit strategy instruction to the curriculum could make learners more confident and perform better under classroom and test conditions.

Although the study contributes importantly to what we know, there are a number of limitations. The sample size was quite modest and included only learners at proficiency levels of B1–B2, which resulted in limited generalizability to other proficiency levels or

settings. The data collection occurred in one cultural and institutional setting during which time learners' perceptions and behaviour might have been affected by this setting.

Additionally, anxiety was only assessed prior and after one speaking session and therefore did not provide any possibility of examining changes over time. Both qualitative and quantitative data were also collected with the use of self-report instruments such as those of learners' strategies in managing anxiety but are prone to bias.

The future studies can overcome the limitations with more expansive and heterogeneous samples with regard to the level of proficiency, age, and test condition. The longitudinal techniques may investigate the anxiety and coping patterns at longer intervals and various speaking tests. The intervention experiments may also investigate other variables of the examiner (e.g., personality, communicating style, familiarity with the learner) as speaking proficiency influencers. The integration of the physiological measures of anxiety (e.g., heart rate, galvanic skin response) with the self-report measures may provide more accurate affect information. The intervention experiments may also investigate the measures of decreasing the anxiety and enhancing the test accomplishment under various test types as well as with various sets of the learner population.

Overall, this study substantiates the argument that foreign language speaking is determined by something other than linguistic ability but by affect, interpersonal ability, and contextual variables as well. Knowing and keeping in mind the variables can convince examiners and teachers to create more fair, favorable, and valid speaking exams beneficial to test-taking on the part of the test-taker and to eventual language acquisition.

Unexpected Findings

The results of this study produced several findings that were not entirely anticipated. First, despite learners becoming more familiar with the assessment process and reporting the use of various self-regulation strategies, their anxiety levels did not significantly decrease from pre-test to post-test. This contrasts with the expectation, grounded in studies such as Cheng, Horwitz, and Schallert (1999), that repeated exposure to speaking tasks, even over a short period, could lower anxiety. The absence of a measurable reduction suggests that brief encounters with tasks, particularly those involving personally sensitive or emotionally charged topics, may be insufficient to influence affective states meaningfully. This finding

also reinforces Horwitz's (2001) and Gregersen and Horwitz's (2002) assertion that language anxiety is a relatively stable construct, requiring sustained intervention to produce notable change.

Second, the influence of examiner gender on performance proved more pronounced than anticipated. Learners assessed by male examiners exhibited greater variation in performance between gender-neutral (GN) and gender-biased (GB) topics than those assessed by female examiners. This result aligns with Lumley and Sullivan's (2005) observation that interlocutor characteristics can subtly influence performance, yet it adds a cultural dimension. In contexts where gender roles are socially salient, as Norton and Pavlenko (2004) noted, the perceived power dynamics between examiner and examinee may shape performance in not purely linguistic ways. This underscores the importance of examiner training to mitigate potential bias and ensure consistency across candidates.

A further unanticipated observation was that not all learners performed worse on GB topics. For some, such topics appeared to enhance engagement and motivation, even when accompanied by elevated anxiety. This challenges the assumption that emotional sensitivity is uniformly detrimental to performance (Yan & Horwitz, 2008). Instead, it suggests that specific learners may channel the heightened arousal associated with controversial or personally relevant topics into greater focus and effort. However, the present study also found that the use of emotional regulation strategies—such as deep breathing or positive self-talk—was not equally effective for all learners, especially for those experiencing high pre-test anxiety. This finding diverges from MacIntyre and Gregersen’s (2012) view that such strategies are broadly effective, indicating that strategy efficacy may vary according to individual differences in personality, prior experience, or emotional resilience.

These unexpected findings highlight that speaking performance in EFL contexts is highly individualised and shaped by complex interactions between cognitive, affective, and social variables. They point to the need for more personalised speaking assessment and anxiety management approaches, particularly in culturally diverse learning environments.

Implications

Pedagogical Implications

These results have some important applications in relation to English speaking teachers, test constructors, and speaking assessors. The notable difference in performance across gender-neutral (GN) as opposed to gender-biased (GB) topics underlines the necessity of being careful in topic choice in speaking lessons and test conditions. There is the possibility of fairness (O'Sullivan, 2002; Kim, 2009) and validity using topics that are affectively balanced, culturally inclusive, and typical of the learners from diverse backgrounds. If the topics get construed as personal, controversial, or biased, the learners run out of fluency and lucidity but fail in linguistic knowledge (Johnson & Tyler, 1998) instead. The application of neutral topics allows more equal opportunities of all the learners exhibiting language proficiency.

A second pedagogical consequence involves foreign language anxiety exerting its ongoing influence on speaking performance. As with Horwitz, Horwitz, and Cope (1986) and Woodrow (2006), this study verifies that even competent users may incur diminished performance when activated by subject content or interpersonal issues during the task at hand. Teachers might include explicit strategies for emotional control—in the form of positive self-talk, mental rehearsal, deep breathing exercises, and visualizing success—in regular classroom activity (Oxford, 1990; MacIntyre & Gregersen, 2012). As a regular feature of daily instruction rather than pre-exam cramming exercises, incorporating such strategies might enhance learner resilience and self-efficacy gradually over time.

The result of this study that examiner sex affected performance, especially with respect to GB issues, requires increased examiner sensitivity as well. Studies on language tests have found that test-taker reactions are affected subtly by interlocutor personal characteristics, such as gender (Lumley & Sullivan, 2005; O'Loughlin, 2002). There are indications that female examiners might unconsciously provide a more encouraging setting; however, all examiners, whether male or female, should be trained to deliver neutral, encouraging, and consistent interactive styles to ensure fair performance outcomes.

Pre-production exposure will also be beneficial in helping the learners acquaint themselves ahead of time with the formal speaking tasks as well. Allowing little time during preparation or informal pre-task discussion can decrease tension and facilitate more

spontaneity in response (Cheng, Horwitz, & Schallert, 1999). Similarly, inclusion of the group activities like pair or small group discussion in the regular lessons can offer low-stakes practice at rehearsing the ideas and getting accustomed with the various topics.

Finally, they align with learner-centered pedagogic approaches with a focus on meeting the language and affective needs of learners. Preparing the learners to be involved in the selection of topics, examining their affective response, and exchanging strategies for managing anxiety can give their sense of agency some support and encourage communicative competency and confidence.

Assessment Implications

From an assessment perspective, the findings have significant bearings on the development, administration, and marking of EFL speaking tasks. The topic should be culturally inclusive and affectively neutral so as not to disadvantage certain groups of test-takers (Iwashita, McNamara, & Elder, 2008). Pre-testing of topics using a large and heterogeneous sample may lessen perceived lack of fairness and irrelevance. The incorporation of neutral or unbiased material may lessen affective interference and invalidity of construct.

Test design should also take into consideration the long-established correlation of performance with anxiety (MacIntyre & Gardner, 1994; Yan & Horwitz, 2008). Through inclusion of warm-up exercises, informal small practices, or informal conversations before the administration of formal tests, levels of anxiety can be brought down and more satisfactory display of competence permitted. In the case of high stakes, provision of loose timing or breaks during activities may be helpful for anxious learners too.

Another aspect that is crucial is that of examiner training. As we did find a significant effect of examiner gender in this study, the training needs to take account of how interactional style, vocal tone, and body language may affect the test-taker's performance (Swain, 2001; Lumley & Sullivan, 2005). It is crucial that there is consistency in the examiners in order to offer fairness, as well as that there be routine monitoring with calibration meetings so that there is reliability in the scoring.

Finally, assuming the entire-systems thinking in the area of performing measurement oftentimes offers deeper insights. Merging quantative measurements with qualitative judgments—in modes of hesitation, physical action, and investment; as a sample—a more holistic Communicative Competence profile may be gained (Bachman & Palmer, 1996). The

incorporation of the learner self-reflection of the performance may be particularly informative in the context of anxiety or sensitiveness to contents which may have distorted outcomes.

Theoretical Implications

The findings contribute to several theoretical perspectives in SLA, task-based learning, and language testing. They reinforce Krashen's (1982) Affective Filter Hypothesis by demonstrating that emotional factors, particularly anxiety, can impede language production, and that the extent of this inhibition varies according to task characteristics. Familiarity, perceived relevance, and emotional sensitivity all appear to influence the height of the affective filter.

The results also extend Skehan's (1998) cognitive approach to task-based performance by underscoring the role of emotional and cultural factors in allocating attentional resources. Discomfort or internal conflict from topic content can reduce fluency and disrupt cognitive control, suggesting that affective dimensions should be more explicitly integrated into models predicting task performance.

From a sociocultural standpoint, the findings align with Vygotsky's (1978) view that learning and performance are socially mediated. The influence of examiner gender indicates that interlocutor variables are an important and often overlooked aspect of communicative performance. Learners' output is shaped not only by linguistic proficiency, but also by their perceptions of and interactions with the examiner.

Finally, the results support Bachman and Palmer's (1996) model of communicative language ability, which highlights the role of strategic competence, including emotional regulation, in effective language use. In high-stakes speaking assessments, managing affect can be as crucial as managing linguistic resources, reinforcing the importance of viewing language performance as a dynamic, context-dependent process in which meaning, emotion, and social relationships are intertwined.

Suggestions for Future Research

On the basis of this study's findings and limitations, several avenues for future research are clear. First, longitudinal studies could trace performance and anxiety changes over weeks or months to see if repeated exposure to varying speaking tasks elicits

progressive reductions in anxiety, as Cheng et al. (1999) reported and if this holds for emotionally sensitive as well as neutral topics.

Secondly, subsequent studies should examine more types of speaking content than only differences between GN and GB. Politically sensitive topics, cultural controversy issues, or substantial personal issues might be introduced to give a more comprehensive understanding of what role topic familiarity, emotional salience, and perceived relevance play on performance and anxiety (O'Sullivan, 2002; Kim, 2009).

The other recommendation is shifting profiles of participants. The current study was conducted only with university students with proficiency levels of B1–B2; the incorporation of members with various levels of proficiency, ages, and study situations (also distant-learning- or professional-training-based) would contribute in terms of overall generalisability of findings (Horwitz et al., 1986; Woodrow, 2006).

Future studies could also be enriched with experimental manipulations of the characteristics of the examiner. Although this study concluded that the gender of the examiner was significant, other examiner characteristics, such as tone of voice, style of interaction, and familiarity with the learner, may also affect performance (Lumley & Sullivan, 2005; Swain, 2001). Supplementing the perspectives of the learners with the perspective of the examiner could also illuminate more the affective processes underlying variation in performance.

Another promising area of study is intervention-based study of regulation of anxiety. Looking at whether one strategy (e.g., instruction in mindfulness (Gregersen, MacIntyre, & Meza, 2014), cognitive reappraisal, or systematic preparation) is superior with some kinds of topics may illuminate which is optimum with what kinds of learners.

Concurrent with rising prominence of technology in testing, research on application of computer-based or AI-based examiners is also worthwhile. Such research can examine whether computer-based interlocutors with different perceived genders or interacting styles induce the same performance and anxiety as human examiners and may therefore allow diverse and inclusive test settings (Chapelle & Voss, 2016).

Qualitative approaches could offer greater insight via the graphing of affective engagement and learner motivation states while speaking. Such procedures as stimulated recall interview, narrative histories, or affect-tracking programs could offer more finely grained insights into how learners cope with troublesome material and what affective processes are at the root of language production (MacIntyre & Gregersen, 2012).

Finally, there would be a desire for cross-cultural work so as to examine the responses of respondents to sensitive material as mediated via sociocultural conventions of gender, emotion, and topic appropriacy. Pursuing this line of thought via Norton (2013), language learning is concurrent with relations of power and identity; understanding of the same relations within cultural schemes would offer insight into the extent of the understanding of the interaction of topic sensitivity and examination-induced anxiety with examiner variables within diverse EFL scenarios.

REFERENCES

Alavi Dana, S. M., Khoshnazar, M., Banyarani, S. P., Mokhtari, F. S., Mohammadie, Z. E., Shamsehkohan, A., & Deravi, N. (2023). Anti-cancer potential of Hypericum spp. with focus on Hypericum perforatum: A review of the literature. Current Traditional Medicine, 9(4), 87–104.

Alazemi, A. F. T. (2023). The consequences of task-supported language teaching via Telegram: Enhancing EFL learners' engagement and emotion regulation. Springer Open Education. https://doi.org/10.1186/s40862-023-00220-6

Ali, S., & Rezaei, S. (2022). The impact of test anxiety on English speaking performance: A systematic review. Language Testing in Asia, 12(1), 10–25. https://doi.org/10.1186/s40468-022-00143-9

Altakhaineh, A. R. M. (2019). The effect of the gender and culture of the IELTS examiner on the examinees' performance on the IELTS speaking test in the UAE context. International Journal of Arabic-English Studies, 19(1), 33–52. https://doi.org/10.33806/ijaes2000.19.1.2

Aubrey, S. (2017). Measuring enjoyment in task-based language teaching: Developing the TBLT enjoyment scale. Applied Linguistics, 38(5), 1–23.

Aydın, B. (2001). A study of sources of foreign language classroom anxiety in speaking and writing classes (Unpublished doctoral dissertation). Anadolu University, Eskişehir, Turkey.

Azkarai, A. (2015). Males and females in EFL task-based interaction: Does gender have an impact on LREs? VIAL – Vigo International Journal of Applied Linguistics, 12, 9–35. https://dialnet.unirioja.es/servlet/articulo?codigo=5457407

Bachman, L. F., & Palmer, A. S. (1996). Language testing in practice. Oxford University Press.

Bem, S. (1981). Gender schema theory: A cognitive account of sex typing. https://www.semanticscholar.org/paper/Gender-schema-theory%3A-A-cognitive-account-of-sex-Bem

Block, D. (2014). Social class in applied linguistics. Routledge.

Bouras, N., & Holt, G. (2007). Psychiatric and behavioral disorders in intellectual and developmental disabilities (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139461306

Bozdağ, B. (2021). Examination of university students' fear of negative evaluation and academic dishonesty tendencies. Participatory Educational Research, 8(3), 176–187.

Brown, H. D. (2000). Principles of language learning and teaching (4th ed.). Pearson Education. https://www.scirp.org/reference/referencespapers?referenceid=1893384

Brown, H. D. (2016). Principles of language learning and teaching (6th ed.). Pearson Education.

Buckingham, L. (1997). Examine gender and language test performance. Language Testing, 14(2), 118–134. https://doi.org/10.1177/026553229701400202

Camden. (2020, August 18). Foreign language speaking anxiety. Baguettes and Bicyclettes. https://baguettesandbicyclettes.com/2020/08/03/foreign-language-speaking-anxiety/

Canevello, A. (2020). Gender schema theory. In Springer eBooks (pp. 1741–1743).

Castelfranchi, C. (2014). The theory of social anxiety: A cognitive approach. Cognition & Emotion, 28(2), 183–199. https://doi.org/10.1080/02699931.2013.812543

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. Cognition and Instruction, 8(4), 293–332. https://doi.org/10.1207/s1532690xci0804_2

Chapelle, C. A., & Voss, E. (2016). 20 years of technology and language assessment in Language Learning & Technology. Language Learning & Technology, 20(2), 116–128. https://scholarspace.manoa.hawaii.edu/bitstream/10125/44464/1/20_02_chapellevoss.pdf

Chen, L., & Lee, H. (2023). Sociolinguistic factors affecting oral proficiency assessment: Examiner gender and topic familiarity. Journal of Language Testing and Assessment, 18(2), 120–138. https://doi.org/10.1016/j.jlta.2023.04.005

Cheng, Y. S., Horwitz, E. K., & Schallert, D. L. (1999). Language anxiety: Differentiating writing and speaking components. Language Learning, 49(3), 417–446.

Cheng, Y. S., Horwitz, E. K., & Schallert, D. L. (1999). Language anxiety: Differentiating writing and speaking components. Language Learning, 49(3), 417–446. https://doi.org/10.1111/0023-8333.00095

Chomsky, N. (1965). Aspects of the theory of syntax. MIT Press.

Coates, J. (2015). Women, men, and language. Routledge. https://doi.org/10.4324/9781315645612

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory, and antiracist politics. University of Chicago Legal Forum, 1989(1), 139–167.

Cunill, M. L. (n.d.). Speech anxiety. Pressbooks. https://fscj.pressbooks.pub/publicspeaking/chapter/speech-anxiety

Deckha, M. (2008). Intersectionality and posthumanist visions of equality. Wisconsin Journal of Law, Gender & Society, 23(2), 249–267.

Donate, Á. (2022). Task anxiety, cognition, and performance on oral tasks in L2 Spanish. Journal of Spanish Language Teaching, 9(1), 45–63. https://doi.org/10.1080/23247797.2022.2053024

Dörnyei, Z., & Ryan, S. (2015). The psychology of the language learner revisited. Routledge. https://doi.org/10.4324/9781315779553

Drew, C. (2023, September 10). Gender bias: 15 examples & definition (2025). Helpful Professor. https://helpfulprofessor.com/gender-bias-examples/

Du, X. (2009). The Affective Filter in Second Language Teaching. Asian Social Science, 5(9), 18–22.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. Annual Review of Psychology, 53, 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Elder, C., & Iwashita, N. (2008). Planning for test performance: Does it make a difference? In R. Ellis (Ed.), Planning and task performance in a second language (pp. 219–238). John Benjamins. https://doi.org/10.1075/lllt.11.13eld

Ellis, R. (2003). Task-based language learning and teaching. Oxford University Press.

Ellis, R. (2005). Planning and task performance in a second language. John Benjamins.

Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. Applied Linguistics, 30(4), 474–509. https://doi.org/10.1093/applin/amp042

Ellis, R. (2017). Task-based language teaching: 2003–2016 review. Language Teaching, 50(4), 507–526. https://doi.org/10.1017/S0261444817000092

Ellis, R., Skehan, P., Li, S., Shintani, N., & Lambert, C. (2019). Task-based language teaching: Theory and practice. Cambridge University Press.

Fulcher, G. (2003). Testing second language speaking. Pearson Education.

Garcia, M., Lopez, S., & Chan, A. (2024). Demographic predictors of public speaking anxiety in university students. Current Psychology. Advance online publication. https://doi.org/10.1007/s12144-024-06216-w

Gardner, R. C., & MacIntyre, P. D. (1993). A student's contributions to second-language learning. Part II: Affective variables. Language Teaching, 26(1), 1–11. https://doi.org/10.1017/s0261444800000045

Getie, A. S. (2020). Factors affecting the attitudes of students towards learning English as a foreign language. Cogent Education, 7(1), 1738184. https://doi.org/10.1080/2331186x.2020.1738184

Gkintoni, E. (2025). Challenging cognitive load theory: The role of AI-based interventions in language learning. Frontiers in Education. https://doi.org/10.3389/feduc.2025.11852728

Gregersen, T., & Horwitz, E. K. (2002). Language learning and perfectionism: Anxious and non-anxious language learners' reactions to their own oral performance. The Modern Language Journal, 86(4), 562–570. https://doi.org/10.1111/j.1540-4781.00161

Hancock, A. B., Stone, M. D., Brundage, S. B., & Zeigler, M. T. (2009). Public speaking attitudes: Does curriculum make a difference? Journal of Voice, 24(3), 302–307. https://doi.org/10.1016/j.jvoice.2008.09.007

Horwitz, E. K., & Young, D. J. (Eds.). (1991). Language anxiety: From theory and research to classroom implications. Prentice Hall.

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. The Modern Language Journal, 70(2), 125–132. https://doi.org/10.1111/j.1540-4781.1986.tb05256.x

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. The Modern Language Journal, 70(2), 125–132. https://doi.org/10.1111/j.1540-4781.1986.tb05256.x

Howe, J. (n.d.). Fear of negative and positive evaluation across social evaluative situations (Master's thesis). Eastern Illinois University, The Keep. https://thekeep.eiu.edu/theses/1265/

Hu, J. (2023). The Survey on Students' Emotional State in Oral English Assessment. Research in Applied Linguistics, 11(2), 426–437. Retrieved from

Hughes, A. (2003). Testing for language teachers (2nd ed.). Cambridge University Press.

Hutto, C. (2024, May 8). Gender-neutral terms for the workplace & beyond. InHerSight. https://www.inhersight.com/blog/diversity/gender-neutral-terms

IMD Business School. (2025, June 17). Understanding gender bias: Key issues and strategies for change. IMD Business School for Management and Leadership Courses. https://www.imd.org/blog/management/gender-bias/

Iwashita, N. (2008). Lexical profiles in EAP speaking task performance. Indonesian Journal of English Language Teaching, 4(2), 111–121.

Iwashita, N., McNamara, T., & Elder, C. (2008). Can we predict task difficulty in an oral proficiency test? Exploring the potential of a multiple-method approach. Language Testing, 25(4), 401–425. https://doi.org/10.1177/0265532208094273

Jankowiak, K., Błaszczak, J., & Dziubalska-Kołaczyk, K. (2024). Bilingual speakers are less sensitive to gender stereotypes in their foreign language. Bilingualism: Language and Cognition, 27(1), 135–149. https://doi.org/10.1017/S1366728923000195

Johnson, K. (2017). An introduction to foreign language learning and teaching. Routledge. https://doi.org/10.4324/9781315734675

Johnson, M., & Tyler, A. (1998). Topic bias in oral proficiency interviews: The case of the "telephone call." Language Testing, 15(2), 211–229. https://doi.org/10.1177/026553229801500204

Karimi, F., & Zhaleh, M. (2024). Anxiety and speaking performance: Mediating effects of task type and topic familiarity. TESOL Quarterly, 58(1), 89–107. https://doi.org/10.1002/tesq.3331

Kim, H. J. (2009). Topic effects in EFL speaking assessment: A comparison of task performance across topics. English Teaching, 64(3), 23–49. https://doi.org/10.15858/engtea.64.3.200909.23

Kim, H. R. (2009). Gender differences in L2 oral task performance. English Teaching, 64(1), 51–76.

Kim, S., & Park, J. (2023). Innovative assessment methods in macroeconomics education: A focus on practical skills. Journal of Higher Education Policy and Management, 45(2), 156–170. https://doi.org/10.1080/1360080X.2022.2159259

Kiruthiga, E., & Christopher, G. (2022). The impact of affective factors on English-speaking skills. Theory and Practice in Language Studies, 12(12), 2478–2485. https://doi.org/10.17507/tpls.1212.02

Kormos, J., & Dörnyei, Z. (2004). The interaction of linguistic and motivational variables in second language task performance. Zeitschrift für Interkulturellen Fremdsprachenunterricht, 9(2), 1–19.

Krashen, S. D. (1982). Principles and practice in second language acquisition. Pergamon Press.

Krashen, S. D. (1982). Principles and practice in second language acquisition. Pergamon.

Krashen, S. D. (1985). Principles and Practice in Second Language Acquisition. Pergamon Press.

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

Krashen, S. D. (1986). The input hypothesis: Issues and implications. Longman.

Kutuk, G., Putwain, D. W., Kaye, L. K., & Garrett, B. (2021). Relations between gender stereotyping and foreign language attainment: The mediating role of language learners' anxiety and self-efficacy. British Journal of Educational Psychology, 92(1), 212–235. https://doi.org/10.1111/bjep.12446

Laurin-Barantke, L., Hoyer, J., Fehm, L., & Knappe, S. (2016). Oral but not written test anxiety is related to social anxiety. World Journal of Psychiatry, 6(3), 351–358. https://doi.org/10.5498/wjp.v6.i3.351

Lin, G. H. C. (2008). Pedagogies Proving Krashen's Theory of Affective Filter. ERIC Document Reproduction Service No. ED503681.

Liu, L., Saeed, M. A., Abdelrasheed, N. S. G., Shakibaei, G., & Khafaga, A. F. (2022). Perspectives of EFL learners and teachers on self-efficacy and academic achievement: The role of gender, culture, and learning environment. Frontiers in Psychology, 13, 996736. https://doi.org/10.3389/fpsyg.2022.996736

Liu, X. (2023). The effect of the affective filter hypothesis on college students' English writing teaching and its enlightenment. OALib, 10(9), 1–9. https://doi.org/10.4236/oalib.1110671

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), Handbook of second language acquisition (pp. 413–468). Academic Press.

Loveless, B. (2023, April 14). Cognitive load theory – The definitive guide. Education Corner. https://www.educationcorner.com/cognitive-load-theory/

Lumley, T., & O'Sullivan, B. (2005). The effect of test-taker gender, audience, and topic on task performance in tape-mediated speaking assessment. Language Testing, 22(4), 415–437. https://doi.org/10.1191/0265532205lt314oa

Lumley, T., & Sullivan, J. (2005). The effects of test-taker characteristics and examiner behaviour on ratings in oral assessment. Language Testing, 22(4), 413–437. https://doi.org/10.1191/0265532205lt302oa

Luoma, S. (2004). Assessing speaking. Cambridge University Press.

MacIntyre, P. D. (2017). An overview of language anxiety research and trends in its development. In C. Gkonou, M. Daubney, & J.-M. Dewaele (Eds.), New insights into language anxiety: Theory, research and educational implications (pp. 11–30). Multilingual Matters.

MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. Language Learning, 44(2), 283–305. https://doi.org/10.1111/j.1467-1770.1994.tb01103.x

MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. Language Learning, 44(2), 283–305. https://doi.org/10.1111/j.1467-1770.1994.tb01103.x

MacIntyre, P. D., & Gregersen, T. (2012). Affect: The role of language anxiety and other emotions in language learning. In S. Mercer, S. Ryan, & M. Williams (Eds.),

Psychology for language learning: Insights from research, theory and practice (pp. 103–118). Palgrave Macmillan. https://doi.org/10.1057/9781137032828_8

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. The Modern Language Journal, 82(4), 545–562. https://doi.org/10.1111/j.1540-4781.1998.tb05543.x

Menard-Warwick, J. (2008). The cultural and intercultural identities of transnational English teachers: Two case studies from the Americas. TESOL Quarterly, 42(4), 617–640. https://doi.org/10.1002/j.1545-7249.2008.tb00153.x

Mohtasham, L., & Farnia, M. (2017). English speaking anxiety: A study of the effect of gender on Iranian EFL university students' perceptions. International Journal of Research in English Education, 2(4), 66–79. https://doi.org/10.29252/ijree.2.4.66

Montijo, S. (2022, March 8). Public speaking anxiety: What is it and tips to overcome it. Psych Central. https://psychcentral.com/anxiety/public-speaking-anxiety

Mora, J. C., Treffers-Daller, J., & Alario, F.-X. (2025). Speaking anxiety and task complexity effects on second language speech. Second Language Research, 41(1), 5–28. https://doi.org/10.1177/02676583231177767

Motha, S. (2014). Race, empire, and English language teaching: Creating responsible and ethical anti-racist practice. Teachers College Press.

MSEd, K. C. (2023, March 14). How observational learning affects behavior. Verywell Mind. https://www.verywellmind.com/observational-learning-2795402

Nakamura, T., Phung, L., & Reinders, H. (2021). The role of task design in learners' affective responses: A dynamic perspective. Language Teaching Research, 25(4), 585–608. https://doi.org/10.1177/1362168820903347

Norton, B. (2013). Identity and language learning: Extending the conversation (2nd ed.). Multilingual Matters.

Norton, B., & Pavlenko, A. (2004). Addressing gender in the ESL/EFL classroom. TESOL Quarterly, 38(3), 504–514. https://doi.org/10.2307/3588351

O'Loughlin, K. (2002). The impact of gender in oral proficiency testing. Language Testing, 19(2), 169–192. https://doi.org/10.1191/0265532202lt225oa

O'Loughlin, K. (2002). The impact of gender in oral proficiency testing. Language Testing, 19(2), 169–192. https://doi.org/10.1191/0265532202lt224oa

O'Neill, M. L., & Sorochan, J. (2014). Anxiety. In T. Teo (Ed.), Encyclopedia of critical psychology (pp. 195–197). Springer. https://doi.org/10.1007/978-94-007-0753-5_98

O'Sullivan, B. (2002). Learner acquaintanceship and oral proficiency test pair-task performance. Language Testing, 19(3), 277–295. https://doi.org/10.1191/0265532202lt206oa

O'Sullivan, B. (2002). Learner acquaintanceship and oral proficiency test pair-task performance. Language Testing, 19(3), 277–295. https://doi.org/10.1191/0265532202lt229oa

Okyar, H. (2023). Foreign language speaking anxiety and its link to speaking self-efficacy, fear of negative evaluation, self-perceived proficiency, and gender. Science Insights Education Frontiers, 17(2), 2715–2731. https://doi.org/10.15354/sief.23.or388

Oxford University Press. (2023, February 28). Understanding the impact of task design on learners' willingness to speak. Oxford University Press ELT. https://teachingenglishwithoxford.oup.com/2023/02/28/understanding-the-impact-of-task-design-on-learners-willingness-to-speak/

Oxford, R. (1990). Language learning strategies: What every teacher should know. Newbury House.

Oxford, R. L. (1990). Language learning strategies: What every teacher should know. Heinle & Heinle.

Öztürk, G. (2012). Foreign language speaking anxiety and learner motivation: A case study at a Turkish state university (Master's thesis, Middle East Technical University)

Patel, A., & Kar, B. (2025). Class presentation phobia for English as a foreign language among undergraduate students: A study of EFL learners. International Journal for Multidisciplinary Research, 7(1). https://doi.org/10.36948/ijfmr.2025.v07i01.37271

Pica, T. (1987). Second-language acquisition, social interaction, and the classroom. Applied Linguistics, 8(1), 3–21. https://doi.org/10.1093/applin/8.1.3

Porter, D. (1991). Affective factors in language testing. In C. J. Alderson & B. North (Eds.), Language testing in the 1990s: The communicative legacy (pp. xx–xx). Macmillan.

Qiu, X. (2019). Functions of oral monologic tasks: Effects of topic familiarity on L2 speaking performance. Language Teaching Research, 23(4), 468–489. https://doi.org/10.1177/1362168817723119

Qureshi, S. A. R. H. M. R. D. H. (2020, November 2). Exploring the significance of speaking skill for EFL learners. International Journal of English Language and Literature Studies.

Rafek, M. B., Ramli, N. H. L. B., Iksan, H. B., Harith, N. M., & Abas, A. I. B. C. (2014). Gender and language: Communication apprehension in second language learning. Procedia – Social and Behavioral Sciences, 123, 90–96. https://doi.org/10.1016/j.sbspro.2014.01.1401

Ren, J., Wan Nik, W. S., Wang, S., & Wang, H. (2025). Language Evaluation Anxiety and Self-Appraisal Ability: Their Impact on Students' Oral English Performance. Environment and Social Psychology, 10(7), 1–14. https://doi.org/10.59429/esp.v10i7.3905

Robinson, P. (2001). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.), Cognition and second language instruction (pp. 287–318). Cambridge University Press.

Rojas-Buñay, E. V., & Guaman-Luna, M. M. (2024). The impact of theme-based instruction: Trending topics on speaking skills, in EFL ninth-grade learners from B.E.S. Presidente Velasco. Deleted Journal, 7(1), 145–154. https://doi.org/10.62452/j8dsam59

Rost, M. (2022, December 6). The affective filter hypothesis. Lateral Communications.

Samuelsson, I. P. (2011). Why we should begin early with ESD: The role of early childhood education. International Journal of Early Childhood, 43, 103–118. https://doi.org/10.1007/s13158-011-0034-x

Sapolsky, R. M. (2005). Why zebras don’t get ulcers. Henry Holt.

Schneider, S., Gruber, T., & Hars, A. (2018). The autonomy-enhancing effects of choice on cognitive load and intrinsic motivation. Self and Identity, 17(3), 234–249. https://doi.org/10.1080/15298868.2017.1392067

Sharma, P. (2024, September 17). Speak like a pro: The importance of public speaking skills & effective tips for students. K8 School. https://k8school.com/importance-of-public-speaking-skills-for-students/

Skehan, P. (1998). A cognitive approach to language learning. Oxford University Press.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In P. M. Lightbown & N. Spada (Eds.), How languages are learned (pp. 114–115). Oxford University Press.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. Educational Psychology Review, 22(2), 123–138. https://doi.org/10.1007/s10648-010-9128-5

Sweller, J., Ayres, P., & Kalyuga, S. (2011). Cognitive load theory. Springer Science & Business Media.

Sweller, J., Roussel, S., & Tricot, A. (2022). Cognitive load theory and instructional design for language learning. In J. W. Schwieter & Z. Wen (Eds.), The Cambridge handbook of working memory and language (pp. 859–880). Cambridge University Press. https://doi.org/10.1017/9781108955638.045

Szyszka, M. (2024). Metacognitive awareness of foreign language anxiety in speaking: Learners' perceptions and strategies. System, 118, 103986. https://doi.org/10.1016/j.system.2024.103986

Tantucci, V. (2025). The role of verbal imitation in language testing. Applied Linguistics. Advance online publication. https://doi.org/10.1093/applin/amaf044

Thorpe, L. (2025). Virtual reality for non-native speakers: Reducing public speaking anxiety through immersive learning. Journal of Language Learning Technology, 29(2), 45–62. https://doi.org/10.1080/09588221.2025.1998765

Toale, M. C., & McCroskey, J. C. (2001). Ethnocentrism and trait communication apprehension as predictors of interethnic communication apprehension and use of relational maintenance strategies in interethnic communication. Communication Quarterly, 49(1), 70–83. https://doi.org/10.1080/01463370109385615

Tsegaye, D., & Kassa, A. (2023). Psychological factors and language performance: The role of anxiety and self-confidence in EFL learners' speaking skills. International Journal

of Language and Education, 7(1), 45–63. https://doi.org/10.1080/2331186X.2023.2167890

Ugoretz, J. (n.d.). The pitchman in print: Oral performance art in text and context. CUNY Academic Works. https://academicworks.cuny.edu/mhc_pubs/1/

Vineetha, C. B., & Arun Kumar, A. (2017). Filters in Second Language Learning: Findings from a Six-Year Study. Azim Premji University Publications.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.

Wei, X. (2012). An introduction to conversational interaction and second language acquisition. Foreign Language Teaching and Research Press.

Willis, D., & Willis, J. (2007). Doing task-based teaching. Oxford University Press.

Willis, J., & Willis, D. (1996). Challenge and change in language teaching. Heinemann.

Woodrow, L. (2006). Anxiety and speaking English as a second language. RELC Journal, 37(3), 308–328. https://doi.org/10.1177/0033688206071315

Woodrow, L. (2006). Anxiety and speaking English as a second language. RELC Journal, 37(3), 308–328. https://doi.org/10.1177/0033688206071315

Yan, J. X., & Horwitz, E. K. (2008). Learners' perceptions of how anxiety interacts with personal and instructional factors to influence their achievement in English: A qualitative analysis of EFL learners in China. Language Learning, 58(1), 151–183. https://doi.org/10.1111/j.1467-9922.2007.00437.x

Yan, J. X., & Horwitz, E. K. (2008). Learners’ perceptions of how anxiety interacts with personal and instructional factors to influence their achievement in English: A qualitative analysis of EFL learners in China. Language Learning, 58(1), 151–183. https://doi.org/10.1111/j.1467-9922.2007.00437.x

Yilmaz, T., & De Jong, E. (2024). The role of native speaker peers on language learners’ fear of negative evaluation and language anxiety. Boğaziçi Üniversitesi Eğitim Dergisi, 41(1), 1–17. https://doi.org/10.52597/buje.1389669

Zhang, L. H., & Liu, X. Y. (2023). Research on the practical path of language-assisted rural revitalisation oriented by five goals. Heilongjiang Social Science, 4, 136–141.

Zhang, X. P. (2011). The implication of affective factors in second language acquisition on foreign language teaching. Journal of Jinan Vocational College, 6, 100–102.

Zhang, Y., & Wang, X. (2022). Examining the effects of test anxiety and examiner gender on oral English performance. Applied Linguistics Research, 14(3), 204–220. https://doi.org/10.1080/14708477.2022.2045698

Zulkiflee, M. F. A., & Nimehchisalem, V. (2022). Speaking anxiety among single-gender and co-educational schools students in an online learning context. Open Journal of Modern Linguistics, 12(4), 438–459. https://doi.org/10.4236/ojml.2022.124032

Zhang, Y., & Wang, X. (2022). Examining the effects of test anxiety and examiner gender on oral English performance. Applied Linguistics Research, 14(3), 204–220. https://doi.org/10.1080/14708477.2022.2045698

Zulkiflee, M. F. A., & Nimehchisalem, V. (2022). Speaking anxiety among single-gender and co-educational schools students in an online learning context. Open Journal of Modern Linguistics, 12(4), 438–459. https://doi.org/10.4236/ojml.2022.124032

RESUME