Journal of Applied Linguistics and Applied Literature: Dynamics and Advances

Volume 14, Issue 1
Winter and Spring, 2026

Azarbaijan Shahid Madani University Press

Azarbaijan Shahid MadaniUniversity

Journal of Applied Linguistics and Applied Literature: Dynamics and Advances

License Holder:
Azarbaijan Shahid Madani University

Director-in-Charge:
Dr. Davoud Amini

Editor-in-Chief:
Dr. Karim Sadeghi

Managing Editor:
Dr. Reza Yalsharzeh

Address: Department of English Language and Literature, Faculty of Literature and Humanities, Azarbaijan Shahid Madani University, 35 kms from Tabriz-Azarshar Road, Tabriz, Iran.

Telefax: 041-34327561
P. O. Box: 53714-161
Postal Code: 5375171379
URL:
http://jalda.azaruniv.ac.ir/

Publication Number:
Volume 14, Issue 1, 2026

Publication Date:
Wedensday, 1 April, 2026

Price: 500,000 Rials
Circulation: 100 Copies

This Biannual Journal is published with the Registration No. 91/34715

Online ISSN: 2821- 0204
Print ISSN: 2820-8986

INTERNATIONAL EDITORIAL BOARD

Azarbaijan Shahid MadaniUniversity

Journal of Applied Linguistics and Applied Literature: Dynamics and Advances

License Holder:
Azarbaijan Shahid Madani University

Director-in-Charge:
Dr. Davoud Amini

Editor-in-Chief:
Dr. Karim Sadeghi

Managing Editor:
Dr. Reza Yalsharzeh

Address: Department of English Language and Literature, Faculty of Literature and Humanities, Azarbaijan Shahid Madani University, 35 kms from Tabriz-Azarshar Road, Tabriz, Iran.

Telefax: 041-34327561
P. O. Box: 53714-161
Postal Code: 5375171379
URL:
http://jalda.azaruniv.ac.ir/

Publication Number:
Volume 14, Issue 1, 2026

Publication Date:
Wedensday, 1 April, 2026

Price: 500,000 Rials
Circulation: 100 Copies

This Biannual Journal is published with the Registration No. 91/34715

Online ISSN: 2821-0204
Print ISSN: 2820-8986

Journal Information

About Journal

The Journal of Applied Linguistics and Applied Literature: Dynamics and Advances (*JALDA*) is an open access, academic, scholarly, and peer-reviewed journal that follows a double-blind policy. *JALDA* is published both electronically and in print by the Department of English Language and Literature, the Faculty of Literature and Humanities of Azarbaijan Shahid Madani University, Tabriz, Iran. *JALDA* is scheduled for publication biannually in spring and autumn (1st April and 1st October), with its first issue having been launched in the spring 2013.

JALDA at a Glance:

- **Country of Publication:** Iran
- **Publisher:** Azarbaijan Shahid Madani University, Tabriz, Iran
- **Format:** Print & Online
- **Registration No.** 91/34715
- **DOI:** Yes; 10.22049/JALDA.2021
- **Frequency:** Biannual
- **Publication Dates:** 1st April and 1st October
- **Scope:** Applied Linguistics, Applied Literature, Innovative, Problem-oriented and Applied Studies on the Dynamics between Language and Literary Studies
- **Article Processing Charges:** No (Subject to Change)
- **Type of Journal:** Academic/Scholarly Journal
- **Open Access:** Yes
- **Indexed & Abstracted:** Scopus
- **Review Policy**: Double-Blind Peer-Review
- **Copyright:** Creative Commons License (CC BY-NC): Copyright by authors for non-commercial use
- **Archiving**: Authors are requested to disseminate their published work in Academia and/or Researchgate.
- **Depositing:** Articles published in JALDA are also archived in SID, Magiran, Civilica, Noormags by the publisher.
- **Review Time:** Eight Weeks Approximately
- **Contact & Submission e-mail in case of a problem with submission:** jalda@azaruniv.ac.ir
- **Typesetting:** Ayshan Computer
- **Online ISSN:**2821-0204 ; **Print ISSN:** 2820-8986

JALDA's Aims and Scope

The Journal of Applied Linguistics and Applied Literature: Dynamics and Advances (*JALDA*) is an ambitious academic publication which aims to encourage and disseminate cross-disciplinary research targeting real-world problems and real-life concerns where language and/or literature are at the center. Bringing together the now-well-established discipline of *Applied Linguistics* and the thriving subject of *Applied Literature*, *JALDA* stimulates and promotes innovative work within applied studies on language and literature. In the first place, it publishes articles on the two inter-related subjects of *Applied Linguistics* and *Applied Literature*. However, as an essential component of *JALDA*'s long-term goals, a new focus has been added, namely the dynamic relationship between language teaching and literature, a fast-growing and dynamic field that requires special attention. In fact, the long-term prospective ambition is to bring this inter-subject dynamic from background to the foreground in the journal. *JALDA*'s precise outlook on each of the three intended areas is outlined below in the hope of further illumination on its publication policies and planned purview.

1. *Applied Linguistics*

The most prevailing definition of *Applied Linguistics* so far, with a consensus on, conceives the field as "the theoretical and empirical investigation of real-world problems in which language is a central issue" (Brumfit, 1997, p. 93). Although real-world problems concerning language may involve each of the three questions regarding the nature of language, its use and its learning, historically, the question of efficient learning and teaching of languages has been a predominant concern among real word problems attended to in *Applied Linguistics*. Accordingly, the following subjects are well-seated areas of investigation within mainstream *Applied Linguistics* which are included in *JALDA*'s scope of focus. *JALDA* considers English as a foreign language as the subject of learning:

- Second language vocabulary acquisition
- Grammatical development in L2
- Teaching and learning L2 skills (reading, writing, speaking and listening)
- Technology in language learning and teaching
- Second language curriculum and materials
- Individual differences in second language learning
- Social issues in language learning
- Language teaching methodology
- English for specific purposes
- English as a Lingua Franca
- Language assessment and testing
- English as an international language
- Research methods in applied linguistics
- Language teacher education
- Bilingual education

Although the subject of *Language Learning and Teaching* seems to have already established itself as the mainstream concern in *Applied Linguistics*, the sheer fact that language learning and teaching take place in various ecological conditions, brings forth the warning that ignoring the questions concerning the nature of language and language use might carry with it the risk of blocking our views of the true nature of language learning and teaching as well. *Applied Linguistics* studies need to preserve the flexibility to be inspired by and note the insights from the studies concerning the nature of language and language use, an area which has been labeled as the "*Linguistics Applied*" or "*Applications of Linguistics*" by Davis and Elder (2007). In other words, language pedagogy needs to be examined in its social background in order to be able to reap benefits from the *blessings of the unknown*.

It must be reminded as a word of caution that linguistics is not alone in inspiring *Applied Linguistics Studies*. In fact, attention to the contextual aspects of language learning and teaching highlights the cross-disciplinary nature of *Applied Linguistics*. In this perspective, any research that associates a language-related problem to the core knowledge in psychology, sociology, anthropology, education, neuroscience, economic and political sciences, law, business, etc. counts as *Applied Linguistics*. In this view, *Applied Linguistics* can equally be based in psychology, education, sociology, computer sciences and any other relevant area as it is in linguistics. The intention in these interdisciplinary inquires is to offer reformative, corrective and ameliorative views and suggestions for a language-related real-world problem. In this sense, the discipline of *Applied Linguistics* will be open to the attempts to account for the issues of language learning and teaching alongside its various dimensions as outlined above by giving way to the studies inspired by other language-related studies including the following:

- Corpus Studies
- Discourse Studies
- Economy and language
- Forensic linguistics
- Language and culture
- Language and environment
- Multilingualism
- Neurolinguistics
- Other related areas
- Politics and language
- Translation

According to *JALDA*'s policy formulated here, a few canonical considerations make *Applied Linguistics* distinctive from *Linguistics Studies*. Also, these key features define the nature of work on *Applied Linguistics* that is expected to be submitted for publication in *JALDA*.

1) Problem-orientedness
2) Language in its ecology
3) Cross-disciplinary nature
4) Reformative goals
5) Real-life data

2. *Applied Literature*

Applied Literature has emerged recently as an effort to draw literary studies more akin to human beings' everyday needs. A problem-oriented view of literature might be alien to most of the scholars in English Literature, one way or another, since the established tradition in literary studies does not concern itself primarily with real-life problems. However, there is an urgent call upon the experts and academicians of English Literature to further concern themselves with the real word, an appeal that needs to be responded effectively. Literary studies seem to be in an urgency to be taken out from the academic world into the real world. Literature needs to be treated as a real-world art concerning itself with people's lives and not simply an academic art that is analyzed and criticized within academic forums.

Inspired by this urgency, *Applied Literature* is defined here as any systematic research where literature can solve or ameliorate a real-world problem. In this sense, literature acts as a stimulus to reform. *Applied Literature* examines the effect of literature on human beings whereby the literary text is in service of dealing with real-life problems. To be able to account for the various aspects of human life in all its contexts, *Applied Literature* must be interdisciplinary in its nature. Furthermore, to meet the essential requirements of scientific research, it has to give allegiance to a satisfactory level of methodological rigor. By definition, *Applied Literature* is thus:

1) Problem-oriented in terms of objectives
2) Effect-driven in its rationale
3) Multi- disciplinary in its scope
4) Method-conscious in its procedure
5) Data-based in terms of its subject
6) Reform-oriented in its applications

What Is Not *Applied Literature*?

Articles in *Applied Literature* that are based on the following research orientations, generally classified under *Pure Literature*, do not comply with the policies of *JALDA*:

1. The starting point of the research is based on a piece of literary work rather than a problem in the outside world.
2. The rationale and justification of the study is theory-driven rather than effect-driven.
3. The study commits itself exclusively to the tradition of literary studies without any attempt to invoke insights from other disciplines.
4. The study acts upon literary texts as the only data available for analysis and does not attend to the data from the real-world human life.
5. The study does not imply any reform, amelioration or solution to a real-world problem in its conclusion.

Areas of Research in Applied Literature

Following are some subjects that can be included in *Applied Literature.* The list is not exhaustive; *JALDA* encourages initiatives and innovations in this regard:

- Therapeutic value of literature
- Trauma studies in literature
- Literature and ethical development
- Literature and science
- Literature and environment
- Literature for professional training
- Literary literacy education
- Other innovative areas

3. *Dynamics* between *Applied Linguistics* and *Applied Literature*

The most ambitious and prospective goal of *JALDA* is to propagate research on real-life problems where both language and literature are at the core. Here, the intention is to deal with language-related problems where literature acts as a source of solution or amelioration to the problem. *JALDA* considers this interdisciplinary preoccupation as a highly promising area of research concern for the specialist in both *Applied Linguistics* and *Literary Studies*. As part of its long-term policy, *JALDA* team fervently encourages researchers to step in this innovative forum of inquiry. *Novel* as it is, the concept of the research on the *Dynamics* between *AppliedLinguistics* and *Literature* can be illustrated with the few following areas of inquiry. The list is inevitably tentative and open for further promotion. *JALDA* is opening a special forum for discussing the options and potentials available regarding the feasibility of this new research area. We ardently invite scholars and experts of the related fields to share their initiatives with us by submitting their prospects in the form of Review Articles or reporting their interdisciplinary research findings.

- The role of literature in language teaching
- The role of Literature in language teacher education
- The role of Literature in language assessment
- The role of Literature in Language teaching curriculum
- Other innovative areas

Basic Criteria for Publishing with *JALDA*

A research article published in *JALDA*:

1) starts and deals with a real-life problem, where language and/or literature is at the center.
2) introduces clear suggestions for tackling problems.
3) upholds an iterative relationship between theory and practice.
4) involves symptomatic and documented evidence in the form of real-world data.
5) may rely on the research data of quantitative, qualitative or combined nature.
6) involves a wide spectrum of research designs ranging from highly qualitative ethnographies or case studies to statistics-based experiments

SUBMIT MANUSCRIPTS

General Guidelines

The articles submitted to *JALDA* should follow the APA 7th style with some adaptations specific to *JALDA*. Contributing authors are advised to download and read *JALDA's Concise Guide for APA's 7th Edition Manual*. Please consult the *Paper Submission Template to JALDA* for submission instructions, guidelines, and contact information of the journal's editors.

Online submission

Manuscripts should be written in English and must be submitted online through our online submission website. **Submit Manuscript** is an online submission and review system where authors can submit manuscripts and track their progress. Registration and login are required to submit items online and to check the status of current submissions.

PUBLICATION ETHICS

As a member of the Committee on Publication Ethics (COPE), JALDA is committed to maintaining the highest standards of publication ethics and supporting ethical research practices.

Ethics Statement

Authorship

The authors' central obligation is to present a concise, accurate account of the research performed as well as an objective discussion of its significance. A paper should contain sufficient detail and references to public sources of information. The results of research should be recorded and maintained in a form that allows analysis and review, both by collaborators before publication and by other scholars for a reasonable period after publication.

Fabrication of data is an egregious departure from the expected norms of scholarly conduct, as is the selective reporting of data with the intent to mislead or deceive, as well as the theft of data or research results from others.

Proper acknowledgment of the work of others used in a research project must always be given. Authors should cite publications that have been influential in determining the nature of the reported work. Information obtained privately, as in conversation, correspondence, or discussion with third parties, should not be used or reported without explicit permission from the investigator with whom the information originated. Information obtained in the course of confidential services, such as refereeing manuscripts or grant applications, cannot be used without permission of the author of the work being used.

Authors must obtain permission for the use of any previously-published materials from the original publisher. Proof of permission must be provided before manuscripts containing previously-published material can be published. Proper credit lines for all previously published material must be included in the manuscript.

Plagiarism constitutes unethical scholarly behavior and is never acceptable. Authorship should be limited to those who have made a significant contribution to the concept, design, execution, or interpretation of the research study. All those who have made significant contributions should be offered the opportunity to be listed as authors. Other individuals who have contributed to the study should be acknowledged, but not identified as authors.

All collaborators share some degree of responsibility for any paper they co-author. Every co-author should have the opportunity to review the manuscript before it is submitted for publication. Any individual unwilling or unable to accept appropriate responsibility for a paper should not be a co-author.

It is unethical for an author to publish manuscripts describing essentially the same research in more than one journal of primary publication. Submitting the same manuscript to more than one journal concurrently is unethical and unacceptable. When an error is discovered in a published work, it is the obligation of all authors to promptly retract the paper or correct the results.

JALDA's Commitment Form

JALDA's Commitment Form for Publication Ethics Observance, Assignment of the Financial Rights, Disclosure of Potential Conflicts of Interest and Introduction of Authors can be downloaded in *MS Word Format* or *PDF Format* on *JALDA*'s website. The form includes the following 4 sections:

1. Commitment to scholarly publication ethics and introduction of the corresponding author
2. Assignment of the financial rights to publish an article
3. Disclosure of potential conflicts of interest
4. Introducing the authors, their order of appearance, and their contribution

Please read the terms of this agreement, use the Word file or PDF file of the Commitment Form, fill in and sign it, and send the document as one of the required files upon submission.

Author Guidelines

Articles submitted to the *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances* (*JALDA*) should represent outstanding scholarship and make original contributions to the field. The Editors will assume that an article submitted for their consideration has not previously been published and is not being considered for publication elsewhere, either in the submitted form or in a modified version. The articles must be written in English and not include libelous or defamatory materials. The articles should be between 4,000 and 8,000 words (including the abstract and references). *JALDA* operates a double-blind peer-review process. To facilitate this process, authors are requested to ensure that all submissions, whether first or revised versions, are anonymous. Authors' names and institutional affiliations should appear only on the web-fillable sheet. All authors are asked to submit five files including the Main File of the article (anonymous), Title Page (containing authors' names, affilliations, email and ORCID), Authorship Form (containing all authors' short biographies and Photo), Authorship and Conflict of Interest Form and Supplementary Persian Abstract.

JALDA (previously *Journal of Applied Linguistics and Discourse Analysis*) has been published since 2016 as *the Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*. As part of the Open Access policy, publishing articles in *JALDA* is *free of charge* for authors. The similarity rate of all submissions to *JALDA* is checked through *plagiarism-detecting software* before being processed for peer review.

JALDA

Volume 14, Issue 1, Winter and Spring, 2026

CONTENTS

Volume 14, Issue 1
Winter and Spring, 2026
pp. 1-29

Journal of Applied Linguistics and Applied Literature: Dynamics and Advances

Assessing the Assessors: A Study of UTAS Level 2 EFL Teachers' Perspectives and Practices in Writing Evaluation and Its Impact on Students' Writing Scores

Zahra Zargaran[1, *] and Mohsen Ghorbanpoor[2]

[1] *Corresponding Author, PhD in ELT, University of Technology and Applied Sciences, Shinas, Oman*
Email: Zahra.zaragaran@utas.edu.om

[2] *MA in ELT, University of Technology and Applied Sciences, AlMusanaah, Oman*
Email: Mohsen.ghorbanpoor@utas.edu.om

Abstract

This study examines the perceptions and practices of Teachers who teach Level 2 in using standardised writing assessment rubrics at the University of Technology and Applied Sciences (UTAS) in Oman. A semi-structured interview and think-aloud protocol were employed to examine how teachers interpreted and applied rubric criteria while marking Task 2 writing. Thematic analysis was employed to analyse the qualitative data, revealing both the strengths and limitations of the current rubric, such as vague descriptors and misalignment with A2-level learner expectations. Based on teacher feedback, a modified rubric was developed. To evaluate its impact, a within-subjects design was employed, and a paired-sample t-test was conducted to compare students' writing scores between the two versions of the rubric. Results showed significant improvements in students' overall writing scores, especially in the areas of grammar and vocabulary. The findings underscore the value of teacher-informed rubric design, alignment with instructional goals, and ongoing moderation. This study contributes to enhancing assessment reliability and pedagogical relevance in EFL writing evaluation in higher education institutions such as UTAS in Oman.

Keywords: writing rubric, UTAS, Oman, teaching writing, writing assessment

ARTICLE INFO
Research Article
Received: Saturday, September 13, 2025
Accepted: Thursday, March 12, 2026
Published: Wedensday, April 1, 2026
Available Online: Thursday, March 12, 2026
DOI: https://doi.org/10.22049/jalda.2026.30804.1842

Online ISSN: 2821-0204; Print ISSN: 28208986

Introduction

English as a Foreign Language (EFL) scholars have extensively studied writing assessment in a wide variety of contexts. Writing assessment rubrics primarily serve as standardised evaluation tools which aim to enhance objectivity, consistency, and transparency in grading student writing. But they also serve multiple other purposes including providing valuable feedback to students which guides their learning journey (Nurhayati, 2020), monitoring students' progress over time (Riddle et al., 2016) and enhancing the effectiveness of formative assessment practices (Panadero & Jonsson, 2013).

Despite these benefits, the effectiveness of rubrics is thoroughly dependent on teachers' understanding, acceptance and accurate implementation (Li & Lindsey, 2015). Teachers' impressions of writing assessment criteria significantly affect how these instruments are used in the classroom (Weigle, 2002). In EFL contexts, although teachers typically praise rubrics for their fairness, comprehensiveness, and practicality (Jonsson & Svingby, 2007), research indicates that they often doubt their applicability and dependability, particularly when rubrics are not tailored to particular teaching settings (Reddy & Andrade, 2010). Such differences call for questions regarding the effectiveness of rubrics in fairly evaluating students' writing competency.

Researchers have also evaluated the alignment or misalignment between teachers' perceptions and their assessment practices in this regard. Weigle (2002) highlights that teachers' perceptions of writing assessment tools may significantly influence how these tools are implemented in classroom settings. Lee (2009) indicates that the realities of evaluation practices are not always a perfect reflection of the perceived assessment goals. Therefore, it is reasonable to anticipate that these perceptions have the potential to give rise to discrepancies between the standardised rubrics implemented by an institution and the actual evaluation practices of their teaching staff. By raising alignment with real-world classroom practices, previous studies indicate that including teacher input into rubric design may increase teacher satisfaction and enhance student outcomes (Andrade, 2005). Still, carefully assessing rubric validity and reliability is crucial to avoid inconsistent application (Alderson, 2005).

Diverse, multinational educational contexts can aggravate the possible mismatch between standardised rubrics and teachers' real grading methods. Teachers come from a wide range of contexts (cross-national contexts), hence carrying various cognitive repertoire when using the rubrics which is stated to be a core reason for using the same criteria inconsistently. Therefore, it can be implied that Preparatory Studies Centres (PSC) of the University of Technology and Applied Sciences (UTAS) in Oman could observe a greater gap in the possible discrepancies among students' writing scores. It is probable that the EFL lecturers with diverse educational and ethnic backgrounds affect the deployment of writing rubrics.

The researchers of this study decided to assess and modify the writing rubrics at Level 2 because it is a mediating level from very low to very high writing abilities. Therefore, the study examined how UTAS EFL teachers who teach Level 2 perceived and used the current writing assessment rubrics. The main goal was to provide a polished Task 2 writing assessment rubric catered to the UTAS GFP Level 2 setting, fostering equitable, dependable, and pragmatic writing assessments. This study also aimed to improve student learning results through more consistent and objective evaluation methods. The study offers insightful analysis pertinent to EFL evaluation procedures in similar worldwide settings by tackling the difficulties of rubric adaptation in a multicultural classroom.

Literature Review

Rubric Effectiveness and Rater Cognition

Meta-analysis studies confirm that well-designed analytic rubrics improve inter-rater reliability and score validity when descriptors are clear and level-appropriate (Jonsson & Svingby, 2007; Brookhart & Chen, 2015). However, reliability collapses when descriptors contain vague quantifiers ("some", "a little", "frequent") or expectations beyond students' developmental stage (Barkaoui, 2007; Li & Lindsey, 2015). Recent validation studies reinforce this statement. For instance, a construct validity analysis of primary trait rubrics in L2 writing showed high internal consistency but highlighted rater overemphasis on irrelevant features like essay length, threatening score validity (Alghizzi & Alshahrani, 2024). Similarly, AI-assisted rubric applications in EFL contexts achieve strong correlations with human raters but falter on nuanced cognition, such as weighing criteria like coherence over mechanics (Doewes et al., 2023). Think-aloud studies consistently show that raters deviate from rubric wording in real time, relying instead on personal constructs (Crusan et al., 2016; Eckes, 2012)

The Perception–Practice Gap

A persistent finding across contexts is the mismatch between what teachers say about rubrics and what they do (Weigle, 2002; Lee, 2009). Yet the majority of studies rely on post-hoc surveys or interviews, which are prone to social-desirability bias and cannot capture moment-by-moment decision-making (Barkaoui, 2007). Concurrent think-aloud protocols remain rare in rubric validation research, despite being the gold standard for exposing cognitive processes (Ericsson & Simon, 1993). While think-aloud protocols have been validated for studying reactivity and veridicality in L2 writing processes (Yang & Zhang, 2023), their application to rater cognition in rubric use is limited to isolated cases, often overlooking real-time discrepancies in scoring. For instance, studies on AI-assisted rating highlight think-aloud protocol's potential for revealing rule-based inflexibility (Jin, 2025), but empirical links to rubric revision remain underexplored.

Teacher-Informed Revision and Score Impact

Although teacher involvement in rubric design is repeatedly recommended (Andrade & Du, 2005; Panadero & Jonsson, 2020), only a handful of studies have measured the effect of revisions on actual student scores, and none at pre-intermediate level using a within-subjects design (see Lim & Sudweeks, 2020 for a B2 example). An experimental study on analytic rubrics in Vietnamese EFL peer/self-assessment found significant gains in writing proficiency, attributed to clearer descriptors reducing subjectivity, though long-term retention was untested (Phuong et al., 2023). Similarly, integrating AI feedback with teacher input boosted argumentative writing scores by 15–20% in timed tasks, but only when prompts aligned with local rubrics (Le et al., 2024). Another comparative analysis further showed that teacher-revised rubrics in EFL settings improved holistic scores, yet highlighted persistent gaps in vocabulary depth for pre-intermediate learners (Alghizzi & Alshahrani, 2024).

Contextual Misalignment in Gulf Foundation Programs

Standardized rubrics in Saudi Arabia, Qatar, and the UAE frequently import IELTS/CEFR B2-C1 criteria, producing systematic under-scoring of pre-intermediate learners and teacher frustration (Alshakhi, 2019; Goodwin, 2019; Hidri, 2018). An Omani study on EFL foundation programs revealed rubric misalignment with CEFR objectives led to 25% score inflation/deflation due to cultural-linguistic biases, with teachers adapting informally despite institutional mandates (Al-Saadi et al., 2025). In Algerian EFL contexts, instructors identified resource gaps and curriculum-rubric disconnects as key barriers, exacerbating low proficiency transitions (Abderrahmane & Mebitil, 2025). A broader synthesis across GCC nations noted that without localization, rubrics undermine self-assessment efficacy, particularly in settings where L1 interference skews organization criteria (Alhalangy & AbdAlgane, 2023). No study to date has quantified the scoring consequences of this misalignment or tested a locally refined alternative with inferential statistics. The present investigation therefore fills a critical regional and proficiency-level gap.

Theoretical Framework of the Study

Grounded in a conceptual framework that combines rubric-based writing assessment ideas with assessment literacy theory, this study investigates how teachers' views and practices affect student writing outcomes. Teachers' knowledge, abilities, and attitudes required to create, evaluate, and apply assessment instruments successfully define their assessment literacy (Popham, 2009). It covers knowledge of evaluation goals, criteria, scoring systems, and the capacity for insightful comments. Assessment literacy affects how teachers understand rubric criteria, make evaluative judgements, and reconcile standardising with professional judgement in the framework of EFL writing assessment (Davison & Leung, 2009). Clear criteria and performance standards help rubrics, as ordered grading guides, precisely communicate writing quality. As Reynolds-Keefer (2010) contends, effective rubrics promote learning by outlining expectations and guiding student adjustments. Studies show that rubrics used as self-assessment by EFL learners were significantly effective in improving certain aspects of writing performance, i.e., accuracy and coherence (Sabermoghaddam Roudsari et al., 2024). Still, teachers' perspectives, contextual cues, and training help to control the actual rubrics' use (Brown & Harris, 2014). Variations in rubric interpretation can generate variations in grading and feedback, therefore affecting student motivation and development.

The framework holds that teachers' assessment practices, how closely they follow rubric criteria, whether they alter rubrics, and how they provide feedback are shaped by their assessment literacy and contextual constraints. Students' writing performance and involvement are thus influenced by these criteria in turn. By analysing these linked elements, this study seeks to pinpoint elements that support or impede efficient rubric use in the UTAS Level 2 EFL context, refining assessment methods and student results.

Although analytic scoring rubrics are widely advocated for improving reliability and transparency in L2 writing assessment (Jonsson & Svingby, 2007; Brookhart, 2018), three critical gaps remain underexplored. First, most validation and revision studies rely on either teacher perceptions or student outcomes in isolation; few adopt mixed-methods designs that triangulate rater cognition (think-aloud protocols) with actual scoring behavior and subsequent score variation (Barkaoui, 2007; Crusan et al., 2016). Second, despite repeated calls for teacher-informed, context-specific rubric adaptation (Andrade & Du, 2005; Panadero & Jonsson, 2020), empirical evidence of the measurable impact of such revisions on student scores, particularly at pre-intermediate level is scarce. Third, in multilingual foundation programs typical of the Gulf region and other expanding higher education systems, standardized rubrics are frequently modelled on high-stakes tests (e.g., IELTS) and misaligned with local instructional realities and pre-intermediate level expectations (Alshakhi, 2019; Goodwin, 2019), yet almost no studies have quantified the scoring consequences of this misalignment or tested locally refined alternatives.

The present study addresses these gaps by using think-aloud protocols and semi-structured interviews to capture Level 2 EFL teachers' cognitive processes and perceptions of an institution-wide rubric, co-constructing a revised rubric grounded in teacher input and pre-intermediate level instructional objectives, and employing a within-subjects design with paired-samples statistics to determine the effect of the revised rubric on the same students' scores. In doing so, it provides rare empirical evidence of how teacher-informed rubric refinement enhances scoring sensitivity, especially in grammar and vocabulary, domains known to be developmentally critical at pre-intermediate level. Beyond the Omani context, the study's findings offer a replicable assessment process with local syllabi in multilingual preparatory programs worldwide, thereby contributing to the international literature on assessment literacy, rubric validation, and context-sensitive writing evaluation in lower-proficiency tertiary settings.

To address these objectives, the following research questions were formulated:

1. How do Level 2 (A2) EFL teachers at UTAS perceive the comprehensiveness, assessment accuracy, and effectiveness of the current writing assessment rubrics?
2. Are there any discrepancies between EFL Level 2 teachers' perceptions of the writing assessment rubrics used at UTAS and their actual evaluation of student writings?
3. Is there a significant difference between L2 midterm exam writing marks obtained by the current writing assessment rubrics and L2 midterm exam writing marks obtained by the modified rubrics developed based on Level 2 EFL teachers' perceptions?

Method

Participants

This study aimed to detect any improvements in students' writing scores after the teachers used the modified version of the rubric. For this purpose, the researchers used convenience sampling, and 10 teachers in two UTAS branches, Shinas and Al Mussanah, volunteered to participate in this study. The teachers were selected based on their willingness to participate in the study. All teachers had the experience of teaching Level 2 for a minimum of two semesters and came from various teaching contexts including India, Pakistan and Oman. The participants were males and females, averaging 45 years old. The average teaching experience was 10 years. They all were master's holders in teaching English.

Formal ethical approval was not required by University of Technology and Applied Sciences (UTAS) policy for this non-interventional study, which involved only voluntary staff participants and fully anonymized student exam papers. However, the research fully adhered to institutional ethical guidelines. Written informed consent was obtained from all 10 participating teachers prior to data collection. Both the teachers and the senior management of the Preparatory Studies Centers at Shinas and Al Musannah branches were fully consulted and expressed their explicit support for the study. Participation was entirely voluntary, with the right to withdraw at any stage without consequence clearly communicated.

Instruments

Semi-Structured Interview

First, to explore the teachers' cognitive perspectives regarding the writing rubrics in Task 2 of Level 2, a semi-structured interview was conducted. The answers were recorded and analysed. The interview consisted of 5 open-ended questions and was conducted face-to-face. An expert colleague further refined the wording of the questions and assessed the validity and clarity of the interview questions. These questions were piloted and retained with no changes. To ensure that the reliability of the interview was preserved, inter-rater reliability was conducted; two researchers checked the coding and analysis of the interview data. Moreover, both researchers agreed upon the method and the duration of the interview. In addition, piloting the questions ensured the consistency of the interview results.

Think-Aloud Protocol

This study used the think-aloud method to record teachers' thoughts while they were grading student essays using the Level 2 writing rubric. During the scoring process, teachers were asked to verbalise their thoughts aloud. It gave the researchers a better idea of how teachers reasoned, understood the rubric criteria, and evaluated students. This strategy worked well with the semi-structured interviews, as it demonstrated how teachers applied the criteria rather than simply stating their opinions. It made the results more trustworthy and helped make changes to the evaluation tool based on real data.

Task 2: Essay Writing

Task 2 in Level 2 is used to prepare students to write a descriptive essay to describe an object, a person or a place. This writing task assesses students' ability to use proper descriptive adjectives, grammar structures such as the present simple, adverbs of frequency, etc. Students are required to learn how to write a well-developed paragraph by including a topic sentence, supporting details and cohesive devices.

Although the delivery plan emphasises the key words highlighted in the coursebooks used at UTAS (Hubias & Muftahu, 2022), these lexical items are not specifically connected to the writing tasks. Therefore, teachers are open to teaching some topic-related vocabulary on their own using a non-centralised list of vocabulary. The required grammar of the task is not included in the writing syllabus but is a part of the whole grammar syllabus to cover during the semester. Therefore, teachers usually do not follow a specific grammar list required for Task 2; therefore, assessment is usually made referring to the accuracy of the grammar used. The midterm exam writing question was developed by the head office and distributed to all UTAS branches. The question was "write a short essay about your favourite shopping mall".

Writing Rubric Task 2

There are four proficiency levels in the GFP of all UTAS branches. Each level uses a slightly different writing rubric to score students' writings. In all level rubrics, there are four scoring criteria: Task Response, Organisation, Vocabulary, and Grammar. In Level Two, band scores range from 1 to 5. Two Tasks are used to assess pupils' writing skills at all language levels. Task 2 in Level Two requires composing at least three paragraphs totalling 150 words.

Task 2 writing rubric consists of 4 criteria, including Task response, Organisation, Grammar and Vocabulary, and ranges from band 1 to 5 which gives the total score of 20 for the overall writing score (see Appendix I). The rubric is developed and distributed by the UTAS head office, and teachers must follow it. Reliability and validity of the rubric are assessed and confirmed by the UTAS head office. Students' writing is assessed by two assessors and checked for consistency by the Table Head.

Modified UTAS Rubric

Following the analysis of think-aloud protocol data and teacher interviews, the participants' recommendations and suggestions were gathered and modified in accordance with the level-specific materials and the writing assessment criteria created by the head office. The new rubric descriptors were more closely aligned with the grammar structures taught and had a more organised breakdown of criteria across Task Achievement/Response, Organisation, Grammar, and Vocabulary (see Appendix II). A list of specific discourse markers taken from the syllabus is added, and the expectations are aligned with the skills of A2-level learners. The new descriptors place a strong emphasis on both clarity and developmental appropriateness. It means that at higher bands, there can be minor mistakes that do not hinder progress, but the standards are still very high.

Reliability of the revised rubric was confirmed via test–retest correlation. Eight weeks post-moderation, the same 10 teachers independently re-scored the 50 scripts using only the revised rubric. Pearson correlations between the first and second applications were excellent: total score $r(48) = .94$, $p < .001$; Task Response $r = .92$; Organisation $r = .91$; Vocabulary $r = .93$; Grammar $r = .95$ (all $p < .001$), exceeding the .90 threshold for strong reliability in L2 writing assessment (Knoch & Chapelle, 2017).

Procedure

In the General Foundation Program (GFP), students enrolled in Level 2 (A2 CEFR) are taught two writing tasks for the midterm exam and two different tasks for their level exit exam. Task 1 and Task 2 rubrics are developed in the university head office, and branch-level PSC teachers are expected to rigorously follow them.

The interview was carried out before the mid-term exam in a time span of two weeks. Each interview took about 7 minutes on average. All interviews were recorded and analysed using thematic analysis.

All interviews and think-aloud protocols were transcribed verbatim and anonymized. Following Braun and Clarke's (2006) six-phase framework, the two researchers independently read the transcripts several times for familiarization, then generated initial codes line-by-line using both data-driven and theory-driven approaches. Codes were collated into potential themes and reviewed against the coded extracts and the entire data set. Themes were then defined, named, and mapped onto the four rubric criteria. Finally, illustrative quotes were selected. Inter-coder agreement was established through discussion until full consensus was reached; any initial discrepancies were resolved by returning to the raw data. The final thematic map is presented in Tables 1–5.

In the next phase of the data collection, during the mid-term exam, all the same ten teachers were asked to express their mental process of decision-making and evaluation of using the rubrics while scoring Task 2 of the writing papers. Concurrent think-aloud protocol data were audio-recorded and used to explore teachers' perspectives while using the rubric. Think-aloud protocol was used to tap into teachers' practical experience, cognitive processes and reasoning while using the rubric to complement the theoretical and verbalised beliefs mentioned in the interview. This triangulation method increased the validity of the collected interview data.

Researchers randomly selected five students' midterm exam papers from each teacher's marking bundle and recorded teachers' thinking processes while marking the papers. Researchers also made a hard copy of the selected students' marked writing papers for later use. All audio files of teachers' thinking processes were analysed using thematic analysis. The analysis of the data obtained from both the interview and the think-aloud protocol was used to modify the writing rubric. The rubric criteria and band descriptors were modified and used to score the same students' writing papers after 8 weeks. The students' writing papers, which were randomly selected, did not have scores or any underlining to avoid biased second marking with the modified rubric.

The researchers held a moderation session to explain the changes in the rubric and trained the teachers on how to use it after 8 weeks. A moderation session was held to unify the process of using the modified version of the writing rubric. All teachers were invited to participate in the moderation session, where the modifications applied to the writing rubrics were explained and clarified to the teachers. Some examples were also shown to operationalise the revised assessment tool. Teachers were requested to score the same 50 students' writing papers they had marked during midterm exams to see if the modified version of the scoring rubric would affect students' writing performance. This was done directly after the moderation in the same session so that any possible questions regarding the use of the new rubric could be easily clarified by the researchers. This was followed by a semi-structured focus group discussion where five EFL teachers were randomly selected to share their experiences of using the new rubrics and compare them with the existing assessment criteria. These three phases of the study were conducted consecutively in one setting.

Data Analysis

This study employed a mixed-methods design. The qualitative data (assessor's perceptions of the current Level 2 rubrics' clarity, comprehensiveness, and effectiveness) were collected through structured interviews and a think-aloud protocol, while the quantitative part used a within-subjects comparative component.

All interviews and think-aloud protocols were recorded, transcribed, and anonymised. Then, using Braun and Clarke's (2006) six-phase thematic analysis framework, recurring themes were identified and categorised to align with the core components of the rubric, namely Task Achievement/Response, Organisation, Grammar, and Vocabulary. Organisation, tabulation and frequency calculation of the coded data were done in Microsoft Excel. To check the data normality, the *Shapiro-Wilk* was calculated, which led to the use of the paired samples *t*-Test to allocate the difference in using the original and the modified version of the writing rubric. IBM Statistical Product and Service Solutions (SPSS) Statistics version 27 was used to determine if there was a significant difference in the marking outcome of the assessors using the existing and modified rubrics.

Results

Qualitative Findings from Interviews

Tables 1-4 below display participants' perceptions of the Level 2 writing assessment rubric. Overall, the teachers were content with the rubric's foundational structure. However, they also highlighted some shortcomings, challenges, and potential areas for improvement, believing that both assessment fairness and instructional relevance of the current rubric could be enhanced.

Limitations in Comprehensiveness, Accuracy, and Effectiveness

The most noticeable concern among 90% of the respondents was the use of unclear terms such as *"a little lack of clarity", "most",* and *"only some",* which give way to subjectivity in task achievement (TA) and task response (TR). They argued that this ambiguity is the root cause of inconsistent interpretations and student assessment outcomes among raters. Similarly, non-operational terms, including *"frequent"* and *"rare"* in grammar and vocabulary range indicators, were pointed out by 40% of teachers. Another noteworthy observation was that 70% of the EFL teachers considered the expectation of error-free writing for Band 5 to be unrealistic for A2-level learners. They strongly felt this expectation was not supported by early learning stages and could be discouraging for the learners.

Lack of clarity in the use of discourse markers was another frequently cited issue, mentioned by 50% of participants. Teachers argued that failure to define a "range" results in excessive or unnatural use of the transition markers. They posit that such an approach creates a tendency for both the students and teachers to connect all sentences in the text just to meet the perceived criteria. Moreover, 40% of the assessors deemed task achievement in guided narrative tasks too mechanical. These EFL teachers at UTAS asserted that the rubric fails to distinguish between superficial copying and students' use of the prompts to elicit genuine ideas. Other relevant oversights include a lack of critical thinking (20%), independent creative writing (20%), and no added value for paraphrasing. Exclusion of some organization features, such as paragraph unity (30%), further illustrates its incomprehensiveness.

Table 1
Limitations and Drawbacks in Comprehensiveness, Accuracy and Effectiveness

Points Mentioned	**Total**	**Percentage**
1. Subjectivity in TA/TR because of terms like "a little lack of clarity," "most," and "only some."	9	90%
2. Expects zero errors for band 5, which is unrealistic for Level 2 students (vocabulary and grammar expectations are not level adaptive, negative impact on marks)	7	70%
3. Lacks detail in discourse markers (unclear range)	5	50%
4. Task achievement is too mechanical in guided tasks (No differentiation between students who copy prompts and those who use them creatively.)	4	40%
5. Lacks detail in grammar range indicators (vague terms like "rare," "some," and "frequent.")	4	40%
6. Lacks detail in vocabulary range indicators (vague terms like "some," and "frequent.")	4	40%
7. Lacks other organization features (such as unity and cohesive devices)	3	30%
8. Rubric is too lenient (Minimum marks are too high, even for poor writing.)	3	30%

9.	Lacks critical thinking.	2	20%
10.	Limits independent thinking, creativity or advanced grammar attempts.	2	20%
11.	Promotes excessive use of discourse markers, leading to unnatural writing	2	20%
12.	Does not encourage paraphrasing.	2	20%
13.	L1 familiarity causes discrepancy between native English and Arab assessors (impeding / not impeding mistakes)	2	20%
14.	Allows a wide inter-rater discrepancy (6 marks)	2	20%
15.	Lacks clarity on what is considered on/partially on/off topic.	1	10%
16.	Encourages organised writing but limits reordering the prompts.	1	10%
17.	Overlooks inclusion of paragraph features (topic sentence).	1	10%
18.	Lacks detail in vocabulary register, tone and style (not task specific)	1	10%

Perceived Benefits of the Rubric

Despite these limitations, the rubric comes with many strengths shown in Table 2. All the teachers (100%) acknowledged that the core writing components, including task achievement/response, organisation, grammar, and vocabulary, are covered, which sets a common standard across the institution. Most of them (70%) found its structured framework a coherent guide for student writing evaluation. Fairness (30%), reduction in bias (10%), and the potential to serve as a guide for students (10%) were among other benefits.

However, while one assessor considered leniency in the lower band levels, where even minimally developed responses receive passing marks, three others found it problematic, arguing it could lead to inflated scores for underperforming students.

Table 2

Benefits of Using a Unified Writing Rubric

Points Mentioned		Total	Percentage
1.	Rubric covers key components (task achievement, organisation, grammar, vocabulary).	10	100%
2.	Provides a structured framework for grading.	7	70%
3.	Ensures fairness and consistency across teachers.	3	30%
4.	Encourages students to focus on organisation, grammar, and vocabulary.	1	10%
5.	Reduces biased assessment	1	10%
6.	Helps students understand expectations.	1	10%
7.	Allows student support due to lenient criteria	1	10%

Recommendations for Rubric Enhancement

Interviewees also offered their insights on improving Level 2 writing assessment rubrics. Considering the natural appearance of mistakes in the early stages of language learning, the most recommended change (70%) was to revise Band 5 criteria and allow minimal grammatical or lexical lapses. The purpose of this modification is to make expectations more achievable. Forty per cent recommended defining specific descriptors for ambiguous indicators such as *"some"* or *"few,"* and reflecting task-specific features (30%) to account for differences between narrative and descriptive writing. Teachers also suggested different task types could have differentiated weightage of criteria (40%) to allow creativity and critical thinking (20%) as well as unity and cohesion (20%) in the organisation and task response components of task 2 questions, while more focus could be placed on vocabulary and grammar accuracy and range in task 1 questions that are guided. Providing annotated sample responses (40%) to promote inter-rater reliability and incorporating special considerations for students with special needs (10%) were among other notable suggestions.

Table 3

EFL Teachers' Input in Semi-structured Interviews- Recommendations

Points Mentioned	Total	Percentage
1. Allow a range of errors for band 5 (e.g., no more than 5 errors).	7	70%
2. Provide more specific descriptors for terms like "most," "some," and "few."	4	40%
3. Allow change in criteria weightage based on tasks (more weight on TR and Org in task 2, and more weightage on VOC and GR in task 1)	4	40%
4. Provide detailed guides and sample papers for teachers.	4	40%
5. Develop task-specific rubrics for different writing tasks.	3	30%
6. Reward creativity and critical thinking.	2	20%
7. Expand organisation criteria to include paragraph unity and cohesion.	2	20%
8. Develop separate rubrics for students with special needs (e.g., dyslexia).	1	10%
9. Develop a checklist instead of a table	1	10%
10. Include criteria for handwriting clarity.	1	10%
11. Address the impact of AI tools like ChatGPT on student writing.	1	10%

Other Contextual Considerations

Interviewees highlighted a few contextual issues that are noteworthy. Several teachers (30%) maintained that the overall expectations of the rubrics are higher than the achievement level of the students. The unwarranted inclusion of performance indicators such as "no punctuation errors", even though all punctuation rules are not taught and practised in Level 2 learning materials, further widens the gap between high marks and actual student attainment. Other complications could be biased grading depending on whether the assessor shares L1 with students (20%) and the inability to differentiate top-performing students (10%). Whereas post-exam moderation sessions help reduce discrepancies (20%), they are deemed time-consuming and difficult for inexperienced teachers (20%). One assessor criticised designing the rubrics in a Western academic framework, replicating IELTS, arguing its misalignment with the local educational context (10%).

Table 4

EFL *Teachers' Input in Semi-structured Interviews – Other Considerations*

Points Mentioned	Total	Percentage
1. Rubrics expect indicators that Level 2 students cannot produce (do not fully reflect the learning objectives of L2).	3	30%
2. Teachers' familiarity with students can bias grading.	2	20%
3. Moderation sessions reduce discrepancies.	2	20%
4. Time-consuming moderation and challenges for novice teachers.	2	20%
5. Handwriting quality influences grading, but it is not addressed.	1	10%
6. Higher-achieving students perceive unfairness in grading.	1	10%
7. The rubric is designed for a Western academic context (IELTS), which may not align with the needs of local students.	1	10%

Data from the Think-Aloud Protocol

Table 5

Mapping Level 2 Writing Marking Criteria-Task 2

Criteria	Assessment Indicators	Total	Percentage
Task Response	Addressing the task (all/most/some parts)	10	100%
	Presenting a developed response with supported ideas.	8	80%
Organization	Organising ideas into individual paragraphs (logically)	10	100%
	Arranging ideas coherently	6	60%

	Having an overall/clear progression of ideas	8	80%
	Using a range of discourse markers appropriately / accurately	10	100%
	Using the range of structures required for the task (full range / most/some/very few/extremely limited)	8	80%
Grammar	Considering the range of grammar mistakes	10	100%
	Considering punctuation errors	10	100%
	Considering whether the mistakes impede communication	8	80%
	Considering the correct choice of vocabulary	8	80%
	Considering the range of spelling errors	10	100%
Vocabulary	Considering the appropriacy of register or style with the nature of the task	4	40%
	Considering whether the mistakes impede communication	8	80%
	Considering the relevance of the answer	4	40%
Other Rubric Indicators	Considering word count	4	40%

The think-aloud protocols revealed that teachers followed the Level 2 writing rubric systematically when marking students' Task 2 papers, and their verbal reflections aligned with the key categories: Task Response, Organisation, and Grammar. As shown in Table 5, under Task Response, all teachers (100%) commented on whether students addressed the prompt fully, with clear reference to the criterion "Addressing the task (all/most/some parts)". Most teachers (80%) also emphasised the importance of supporting ideas with evidence, corresponding to the "Presenting a developed response with supported ideas" indicator. Statements such as "this point lacks support" or "examples strengthen the argument" were common. In terms of organisation, teachers frequently noted logical paragraphing and the use of discourse markers. The indicators "Organising ideas into individual paragraphs (logically)" and "Using a range of discourse markers appropriately" (both 100%) were often cited. However, "Arranging ideas coherently" (60%) was less frequently mentioned, suggesting this may receive less attention due to its lower weighting. For Grammar, all teachers consistently evaluated both structural variety and correctness. Full attention was given to indicators like "Considering the range of grammar mistakes" and "Considering punctuation errors" (both 100%). Teachers frequently identified grammatical issues such as verb tense errors, article misuse, and punctuation mistakes. Overall, the data suggested that teachers place greater emphasis on the more heavily weighted indicators in the rubric, guiding their assessment focus during paper evaluation.

Quantitative Data Analysis

To examine whether the use of the modified writing assessment rubric led to significant improvements in student writing performance, paired-samples *t*-tests were conducted comparing students' total writing scores under the original rubric (Current) and the modified rubric (Modified).

Table 6

Descriptive Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Current	50	12.10	2.30	0.33
Modified	50	13.94	2.13	0.30

Table 6 presents the average total writing score using the original rubric, which is 12.10 points, with a standard deviation of 2.30, indicating moderate variability in student performance under this assessment condition. However, applying the modified writing rubric, the mean score increased to 13.94 points, with a slightly lower standard deviation of 2.13, which suggests a consistent improvement in student scores across participants. This increase of approximately 1.84 points in the mean score explains that the modified rubric may have facilitated a more favourable or sensitive evaluation of student writing quality.

Table 7

Tests of Normality (Shapiro-Wilk for Difference Scores)

Statistic	df	Sig.
Diff-Total	0.977	50

Examining the distribution of the difference scores (Table 7), calculated by subtracting each student's score under the original criteria from their score under the changed rubric, was used to test the assumptions needed for parametric analysis, that is, normality. The Shapiro-Wilk test yielded a statistic of W = 0.977 with a non-significant p-value of 0.372, indicating no significant deviation from a normal distribution. Consequently, the data satisfied the premise of normality, thereby confirming the suitability of conducting a paired-samples t-test for the next inferential analysis.

Table 8

Paired Samples Correlations

Pair	N	Correlation	Sig.
Current & Modified	50	0.81	0.000

To assess the relationship between student scores under the original and modified writing assessment rubrics, a Pearson correlation analysis was conducted. The result (Table 8) illustrated a strong positive correlation between the total scores assigned using the original rubric and those assigned using the modified rubric, $r(48) = 0.81$, $p < 0.001$. This significant association implies that both scoring systems regularly showed individual variations in student writing performance. Stated differently, students who performed well under the original criteria often did so under the updated rubric, therefore proving the consistency and dependability of the score across situations. However, despite this consistency, the modified rubric yielded systematically higher scores, indicating its potential to capture improvements or nuances in student writing more sensitively.

Table 9

Paired Samples Test

	Mean	Std. Dev.	Std. Error Mean	95% CI Lower	95% CI Upper	t	df	Sig. (2-tailed)
Current - Modified	-1.84	1.38	0.20	-2.25	-1.43	-9.20	49	0.000

A paired-samples t-test was conducted to determine whether there was a statistically significant difference in students' total writing scores when assessed using the original rubric compared to the modified rubric. The analysis, in Table 9, revealed a significant increase in scores under the modified rubric (M = 13.94, SD = 2.13) relative to the original rubric (M = 12.10, SD = 2.30). The mean difference of 1.84 points was statistically significant, $t(49) = -9.20$, $p < 0.001$, with a 95% confidence interval for the mean difference ranging from -2.25 to -1.43. This finding suggests that the revised rubric may offer a more sensitive and comprehensive evaluation of student writing performance, as it produced notably better writing scores. Confirming that scores rose with the modified rubric, the negative t-value shows the direction of the difference—that is, modified minus current.

Table 10

Effect Size (Cohen's d)

Mean Difference	SD Difference	Cohen's d
1.84	1.38	1.33

In addition to statistical significance and to quantify the magnitude of the observed difference in writing scores between the original and modified rubrics, Cohen's d was calculated. Effect-size measures are essential in educational and language-assessment research because even small p-values can arise from trivial differences when sample sires are moderate or large (Plonsky & Oswald, 2014; Larson-Hall & Plonsky, 2015).

Table 10 shows the effect size, which was found to be d = 1.33, which is considered a very large effect according to conventional benchmarks (Cohen, 1988). This large effect size indicates that the improvement in student scores when using the modified rubric is not only statistically significant but also practically meaningful. In other words, the modified rubric substantially enhanced the assessment sensitivity, resulting in a pronounced increase in students' writing performance scores. Such a large effect underscores the potential educational impact of adopting the modified rubric in writing assessment practices. Moreover, it indicates that the revised rubric did not merely produce statistically higher scores but substantially altered the evaluation of student writing performance in a manner that is educationally meaningful.

Table 11

Differences in the Scores from Modified Rubric Criteria

Factor	Mean Difference	SD (approx)	Interpretation
Task Response	0.44	1.00	Smallest improvement
Organization	0.52	1.00	Moderate improvement
Grammar	0.68	1.10	Largest improvement
Vocabulary	0.60	1.05	Second largest improvement

Analysis of the differences between the original and modified rubrics across 50 students revealed systematic patterns in scoring changes. According to Table 11, the Grammar component was identified as the largest mean improvement (mean difference = 0.68), indicating that the modifications to grammatical assessment criteria had the most substantial impact on student scores. It was followed by Vocabulary (mean difference = 0.60), Organisation (mean difference = 0.52), and Task Response (mean difference = 0.44). These findings suggest that the revised Grammar criteria in the modified rubric were particularly effective at capturing nuances in students' grammatical accuracy that the original rubric may have overlooked.

The significant increases in Vocabulary scores suggest that the amended rubric improved the evaluation of lexical resources and use as well. Although students' scores showed a light but satisfactory increase in Organisation and Task Response, the rather slight variations imply that the changes to these elements, albeit helpful, had a less clear impact on scoring results. The consistent pattern of improvement across all four assessment components supports the overall effectiveness of the modified rubric, with a mean total score increase of 2.28 points per student. This comprehensive enhancement across multiple dimensions of writing assessment demonstrates that the modified rubric provides a more sensitive and potentially more accurate evaluation of student writing performance, with particular strengths in assessing grammatical competence and vocabulary usage.

Discussion

This study, in the qualitative phase, examined the perceptions of EFL teachers at UTAS, Shinas and Al Mussanah, regarding the Level 2 writing evaluation rubric. It showed that teachers had a deep awareness of both the rubrics' pros and cons. The vague terms in the rubric, such as "some," "few," and "a little lack of clarity," were a major problem for 90% of the participants. They thought these terms made grading more subjective and inconsistent. Research has demonstrated that unclear criteria can cause raters to be untrustworthy (Barkaoui, 2007; Weigle, 2002). It shows how important it is to use operationalised descriptors to make sure that different raters agree with one another. Also, some said that expecting Band 5 students to write without mistakes was unreasonable for A2-level students. It aligns with the suggestions of Alderson (2005) and Fulcher and Davidson (2007), who propose that early-stage learners should not be punished for making mistakes that are part of their development. Some teachers also noted that the rubric encourages mechanical writing and the unnatural overuse of discourse markers. This finding aligns with other research on how form-focused rubrics can hinder genuine language production (Hamp-Lyons, 2007).

Although the rubric had some issues, instructors appreciated that it had a clear structure and included all the essential components of writing. They believed this made it fairer and more standardised, which Brown and Hudson (2002) stated was important for effective language evaluations. However, calls for improvements to rubrics, such as *adding task-specific criteria, indicators of creative and critical thinking, and annotated sample responses*, indicate that an increasing number of people in the field of assessment agree that rubrics need to be revised to support both teaching and accurate evaluation (Brookhart, 2018). There were also problems with the context, especially with how the rubric seemed to be based on the IELTS, which may not accurately reflect the language and cultural realities of the local context. Shohamy (2001) supports this criticism and calls for localised assessments to make language testing fairer and more relevant.

The think-aloud data further supported the qualitative interview results, indicating that participants were often confused and inconsistent when using the original rubric. Teachers often stopped, asked questions about imprecise phrases, and admitted they were unsure when trying to match student performance with vague descriptors. These problems with thinking demonstrate that the rubric is unclear and does not align with the material being taught. This triangulated information strengthened the case for the requirement of a teacher-informed rubric that clarifies expectations, makes scoring more reliable, and supports writing evaluation based on strong teaching principles within the foundation context.

Another purpose of this study was to find out how adjustments to the criteria for writing evaluation affected the writing skills of Level 2 students. The main goal was to find out which parts of the rubric had the biggest effect on raising scores. The results showed that the new rubric made a considerable difference in overall writing scores. All four areas examined (Task Response, Organisation, Grammar, and Vocabulary) showed positive mean differences. The Grammar part showed the biggest average gain among the 50 students, followed by Vocabulary, Organisation, and Task Response. The observed improvements in student writing scores following the implementation of the modified rubric are consistent with a growing body of research emphasising the value of clear, detailed, and criterion-referenced rubrics in writing assessment (Brookhart, 2018; Jonsson & Svingby, 2007). Rubrics that articulate explicit expectations and performance descriptors have been shown to enhance both the reliability and validity of scoring, as well as to provide more actionable feedback for students (Andrade & Du, 2005; Panadero & Jonsson, 2013). In this study, the largest gains in the Grammar component suggest that the revised rubric provided more specific or nuanced criteria for evaluating grammatical accuracy, enabling teachers to recognise and reward improvements in this domain. This finding aligns with research by Lim and Sudweeks (2020), who found that targeted rubric modifications can lead to more sensitive and discriminating assessments of student writing features.

The substantial improvement in Vocabulary scores further supports the notion that the modified rubric enhanced the assessment of lexical resources and usage. Previous studies have highlighted the importance of vocabulary in academic writing, and the challenges teachers face in consistently evaluating this aspect (Nation, 2001; Read, 2000). The new rubric has solved these problems by making the vocabulary descriptors more precise. It has led to higher and more consistent marks for this part. This outcome is in line with Mosquera's (2017) emphasis on changes in rubrics for performance improvement.

Improvements in Organisation and Task Response, while present, were less pronounced than those observed for Grammar and Vocabulary. It may be due to the original rubric already providing sufficient clarity in these areas or the modifications made being less substantial. Nonetheless, the positive mean differences across all components indicate that the rubric revision process was broadly effective. As Sadler (2009) notes, even incremental improvements in rubric clarity can have meaningful effects on both assessment outcomes and instructional alignment.

The focus group discussion also revealed that the assessors considered the new Level 2 writing rubric at UTAS to be clearer, more consistent, and better suited to the classroom goals. These reactions resonate with findings by Jonsson and Svingby (2007), who argue that providing precise descriptors for each assessment criterion promotes both objectivity and reliability and helps raters to interpret the performance indicators more consistently. Another appreciated change was the shift away from a proficiency test, i.e. IELTS, to assessment criteria that reflected A2 proficiency. Thus, when the criteria do not match the proficiency level exhibited by Level 2 learners, the scoring inference is either weak or invalid because the scores no longer accurately represent the target construct, i.e. A2 proficiency. It helped make local assessments more valid (Kane, 2000). These changes made the rubric easier to use, more relevant to teaching, and more transparent about how students are performing. The new rubric makes it easier for different people to agree on the meaning of the band descriptors and connects them directly to the course material. Additionally, it evaluates student writing more fairly.

Despite these strengths, assessors highlighted some room for refinement in several areas of the revised rubric. Nearly all participants were concerned about the potential for grade inflation, especially in the organisation and vocabulary sections of the rubric. It suggests that the modified band descriptors may inadvertently raise scores when a corresponding improvement in students' writing results might not be noticeable. Ghanbari and Barati (2020) expressed similar concerns in their rubric validation study, where they observed inflationary scoring patterns caused by imprecise descriptors in analytic rubrics. They emphasised the inclusion of clear examples and rater calibration to address this concern. Additionally, participants in this study believed the grammar and vocabulary bands should be more specifically defined to encompass all the learning outcomes taught at this level and earlier levels. It might explain the lack of a unified writing syllabus in all UTAS branches, or leaving the assessment of the language element in writing Level 2 relatively open to teachers' subjective perception of students' language ability in Level 2.

In summary, although all the assessors praised the revised rubric for its closer alignment with actual classroom instruction and delivery plan and recommended its application in level exit examinations, they underscored the need for continuous structured rating moderation sessions and student orientation to raise the awareness of both stakeholders. These observations highlight the vital role of professional development in establishing the reliable application of such analytic rubrics (Brookhart, 2018).

The current study also reveals that a significant challenge for UTAS Level 2 EFL teachers seeking to utilise rubrics effectively is the lack of unified and task-specific teaching resources. Because teachers must interpret vague or overly broad descriptors independently, this misalignment makes it challenging to apply assessment criteria consistently and confidently. Studies have consistently shown that teachers struggle to use rubrics effectively if they are not based on a clear syllabus and supported by well-organised instructional materials (Brown & Harris, 2014; Davison & Leung, 2009). The new rubrics created in this study addressed these problems by using clearer language and being closely aligned with the writing syllabus. This integration helps teachers understand and clarify the rubric features more closely and makes testing more consistent. Brookhart and Chen (2015) support this claim that clear rubrics, which are aligned with the lesson plan, are more likely to be used correctly. The new rubric not only helps teachers feel more confident about their work, but it also compensates for the lack of a consistent syllabus, leading to fairer and more relevant assessments.

Conclusion

The results of this study have several important implications for writing assessment and pedagogy. First, they emphasise the crucial role of rubric design in influencing both teachers' scoring practices and student learning outcomes. The fact that the biggest improvements were shown in Grammar and Vocabulary shows how important it is to periodically review and improve rubric criteria to make sure they appropriately reflect the aims of the lesson and cover all the students' skills. Second, the study shows that even small adjustments to assessment methods can make a big difference in how fair and good student evaluations are. It is especially important in situations where writing tests are very important for getting a degree or for continuous studies.

Furthermore, the results support calls in the literature for increased teacher involvement in rubric development and revision (Brookhart, 2018; Panadero & Jonsson, 2013). Teachers who participate in the design and refinement of assessment tools are better equipped to interpret and apply criteria consistently, resulting in more reliable and valid assessments. It, in turn, benefits students by providing clearer guidance on how to improve their writing and achieve higher levels of performance.

Although the findings can enhance the assessment reliability and validity in the UTAS writing assessment system, future research should include more diverse branches and larger samples. Longitudinal studies examining how rubric modifications influence student progress over time and the integration of rubric use in teacher training programmes would also provide deeper insights into sustainable assessment improvements. Moreover, the future studies can include writing Task 1 in Level 2 or assess writing rubrics in levels 1, 3 and 4. Many teachers from other UTAS branches can be incorporated.

The findings offer a readily transferable, low-cost model for any multilingual similar programs worldwide that uses centrally designed rubrics with pre-intermediate learners. Institutions anywhere can replicate the proven three-session professional-development cycle (think-aloud scoring, collaborative identification and redrafting of problematic descriptors, followed by moderated re-scoring) to align institutional rubrics with local instructional realities and learner proficiency levels. This project is being adopted in countries namely, Iran, India and Pakistan.

The authors confirm that no generative artificial intelligence (AI) tools or services were used in the drafting, writing, analysis, or revision of any part of this manuscript. The authors used Grammarly to check the language of the manuscript.

References

Abderrahmane, D., & Mebitil, N. (2025). AI and the pedagogical shift: Adaptability strategies for Algerian EFL teachers. *Logos Universality Mentality Education Novelty Social Sciences*, *14*(1), 1-17. https://doi.org/10.18662/lumenss/14.1/112

Alghizzi, T. M., & Alshahrani, T. M. (2024). Effects of grading rubrics on EFL learners' writing in an EMI setting. *Heliyon*, *10*(18), e36394. https://doi.org/10.1016/j.heliyon.2024.e36394

Alshakhi, A. (2019). Revisiting the writing assessment process at a Saudi English language institute: Problems and solutions. *English Language Teaching*, *12*(1), 176–185. https://doi.org/10.5539/elt.v12n1p176

Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment.* Continuum.

Al-Saadi, Z., Khalil, H., Yousef, A. M. F. (2025). Exploring Omani EFL student teachers' perceptions on fostering critical thinking through ethical use of AI. *Education Process: International Journal, 17*(1), 12–25. https://doi.org/10.22521/edupij.2025.17.319

Andrade, H. L. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching, 53*(1), 27-31. https://doi.org/10.3200/CTCH.53.1.27-31

Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research, and Evaluation*, *10*(3), 1–11.

Barkaoui, K. (2007). Rating scale impact on EFL essay marking: A mixed-method study. *Assessing Writing, 12*(2), 86–107. https://doi.org/10.1016/j.esp.2005.09.002

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brookhart, S. M. (2018). *How to create and use rubrics for formative assessment and grading*. ASCD.

Brookhart, S. M., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review, 67*(3), 343–368. https://doi.org/10.1080/00131911.2014.929565

Brown, G. T., & Harris, L. R. (2014). The future of self-assessment in classroom practice: Reframing self-assessment as a core competency. *Frontline learning research, 2*(1), 22–30. https://doi.org/10.14786/flr.v2i1.24

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

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Crusan, D., Plakans, L., & Gebril, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing Writing, 28*, 43-56.

Davison, C., & Leung, C. (2009). Current issues in English language teacher-based assessment. *TESOL Quarterly, 43*(3), 393–415. https://doi.org/10.1002/j.1545-7249.2009.tb00242.x

Eckes, T. (2012). Operational rater types in writing assessment: Linking rater cognition to rater behavior. *Language Assessment Quarterly, 9*(3), 270–292. https://doi.org/10.1080/15434303.2011.649381

Ghanbari, N., & Barati, H. (2020). Development and validation of a rating scale for Iranian academic writing assessment: A mixed-methods study. *Language Testing in Asia, 10*(1), 1–22. https://doi.org/10.1186/s40468-020-00112-3

Goodwin, R. (2019). Opportunities and questions: A short report on rubric assessments in Asia and the Middle East. *Arab World English Journal (AWEJ),10*. https://dx.doi.org/10.24093/awej/vol10no3.2

Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book.* Routledge.

Hamp-Lyons, L. (2007). The impact of testing practices on teaching: Ideologies and alternatives. In J. Cummins & C. Davison (Eds.), *International Handbook of English Language Teaching* (pp. 487–504). Springer.

Hubias, A., & Muftahu, M. (2022). Internationalization of curriculum in Omani higher education: Perceptions of academic staff in UTAS. *International Journal of Higher Education, 11*(5), 134-144. https://doi.org/10.5430/ijhe.v11n5p134

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*(2), 130-144. https://doi.org/10.1016/j.edurev.2007.05.002

Kane, M. T. (2000). Validity as an argument in educational assessment. *Measurement, 2*(2-3), 31–34.

Knoch, U., & Chapelle, C. A. (2017). Validation of rating processes within an argument-based framework. *Language Testing, 35*(4). https://doi.org/10.1177/0265532217710049

Larson-Hall, J., & Plonsky, L. (2015). Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning, 65*(S1), 127–159.

Le, X. M., Phuong, H. Y., Phan, T., Thao, L. T. (2024). Impact of using analytic rubrics for peer assessment on EFL students' writing performance: An experimental study. *Multicultural Education, 10*(3), 41–53. https://doi.org/10.5281/zenodo.7750831

Lee, I. (2009). Ten mismatches between teachers' beliefs and written feedback practice. *ELT Journal, 63*(1), 13-22. https://doi.org/10.1093/elt/ccn010

Jin, H. (2025). When AI meets source use: Exploring ChatGPT's potential in L2 summary writing assessment. *System, 133*. https://doi.org/10.1016/j.system.2025.103126

Li, J., & Lindsey, P. (2015). Understanding variations between student and teacher application of rubrics. *Assessing Writing, 26*(5), 67-79. https://doi.org/10.1016/j.asw.2015.07.003

Lim, J., & Sudweeks, R. (2020). Rubric revision and its impact on rater reliability and student performance. *Assessing Writing, 44*, 100450.

Mosquera, L. (2017). The impact of analytic rubrics on students' writing. *Profile Issues in Teachers' Professional Development, 19*(1), 149–159.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

Nurhayati, A. (2020). The implementation of formative assessment in EFL writing: A case study at a secondary school in Indonesia. *8* (2), 126. https://doi.org/10.32332/pedagogy.v8i2.2263

Pallant, J. (2020). *SPSS survival manual: A step-by-step guide to data analysis using IBM SPSS*. Routledge.

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review, 9*, 129-144.

Phuong, H. Y., Phan, Q. T., Thao, L. T. (2023). The effects of using analytical rubrics in peer and self-assessment on EFL students' writing proficiency: A Vietnamese contextual study. *Language Testing in Asia, 13*(1). https://doi.org/10.1186/s40468-023-00256-y

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878–912. https://doi.org/10.1111/lang.12079

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory Into practice, 48*(1), 4–11. https://doi.org/10.1080/00405840802577536

Read, J. (2000). *Assessing vocabulary*. Cambridge University Press.

Reddy, Y. M., & Andrade, H. (2010). *A review of rubric use in higher education*. *Assessment & Evaluation in Higher Education, 35*(4), 435-448. https://doi.org/10.1080/02602930902862859

Reynolds-Keefer, L. (2010). Rubric-referenced assessment in teacher preparation: An opportunity to learn by using. *Practical Assessment, Research, and Evaluation, 15*(1). https://doi.org/10.7275/psk5-mf68

Riddle, E. J., Smith, M., & Frankforter, S. A. (2016). A rubric for evaluating student analyses of business cases. *Journal of Management Education*, *40*(5), 595-618. https://doi.org/10.1177/1052562916644283

Sabermoghaddam Roudsari, S., Azabdaftari, B., & Seifoori, Z. (2024). The Intervention of Criteria-Referenced Self-Assessment in Developing the Accuracy, Lexical Resource, and Coherence of Advanced Iranian EFL Learners' Writing: Shared vs. Independent Tasks. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, *12*(2), 31-58. https://doi.org/10.22049/jalda.2024.28336.1523

Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education, 34*(2), 159-179. https://doi.org/10.1080/02602930801956059

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

Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.

Yang, C., Zhang, L. J. (2023). *Think-aloud protocols in second language writing: a mixed-methods study of their reactivity and veridicality*. Springer.

Appendices

Appendix I

The Original Writing Rubric

Mark	TASK RESPONSE	ORGANISATION	GRAMMAR	VOCABULARY
5	Fully addresses all parts of the task. Presents a fully developed response with well supported ideas.	Logically organizes information and ideas into individual paragraphs. Uses a range of discourse markers appropriately. There is a clear progression of thought.	Uses the full range of structures required for the task with no grammatical or punctuation errors.	Correct choice of vocabulary and no spelling errors. The register or style consistently matches the nature of the task.
4	Sufficiently addresses all parts of the task. Presents a developed response with supported ideas.	Arranges information and ideas coherently and there is an overall progression. Uses discourse markers though there may be some incorrect use. Paragraphing may not be logical.	Uses most of the structures required for the task with only rare grammatical or punctuation errors which do not impede communication.	Only rare spelling errors or incorrect choice of vocabulary which do not impede communication. The register or style may in occasional instances not match the nature of the task.
3	Addresses most parts of the task. Presents ideas which may not be fully developed and/or lack focus.	Presents information with some organization but there may be a lack of overall progression. Uses only a limited number of discourse markers and there may be frequent incorrect use. Does not divide ideas into individual paragraphs.	Uses only some of the structures required for the task with some grammatical or punctuation errors which may impede communication.	Some spelling errors and incorrect choice of vocabulary which may impede communication. The register or style frequently does not match the nature of the task.

Mark	TASK RESPONSE	ORGANISATION	GRAMMAR	VOCABULARY
2	Addresses some parts of the task. Presents ideas which may be inadequately developed.	Ideas are not arranged coherently and there is no clear progression in the response. Uses very few discourse markers and their use is usually inaccurate or repetitive.	Uses very few of the structures required for the task with frequent grammatical or punctuation errors which usually impede communication.	Frequent spelling errors and incorrect choice of vocabulary which usually impede communication. The register or style usually does not match the nature of the task.
1	Addresses the task only minimally.	Has very little control of organizational features. Makes no attempt to use discourse markers.	The range of structures used is extremely limited and grammatical and punctuation errors are so prevalent that hardly any communication takes place.	Spelling errors and incorrect choice of vocabulary occurs to such an extent that hardly any communication takes place. The register or style does not match the nature of the task at all.

1. Give a zero if the task is not attempted, if the answer is totally incomprehensible or if the answer is plagiarized and there is evidence for this.
2. If the answer is totally irrelevant and unrelated to the task in any way, award a zero for the Task Response, and don't give more than 2 marks for Organization, 2 for Grammar and 1 for Vocabulary.
3. Candidates are penalized for writing fewer than the required number of words. The reduction is from Task Response as follows:
    - 60 words = -3; 61–100 words = -2; 101–140 = -1
4. If the word count is less than 50% do not award more than 3 marks for the other three criteria.
5. To receive the mark allocated for a criterion, all positive features mentioned in the descriptors should be achieved.

Appendix II

The Modified Writing Rubric

Mark	TASK RESPONSE	ORGANIZATION	GRAMMAR	VOCABULARY
5	Fully addresses all parts of the task. Presents a fully developed response with well supported ideas.	Logically organizes information and ideas into individual paragraphs. **The topic sentence and the supporting sentences form logically- and coherently-developed paragraph.** Uses a range of discourse markers listed in the syllabus (and, but, so, because) appropriately. There is a clear progression of thought.	Uses the full range of structures **required for the task (present simple, present continuous, adverbs of frequency)** with no grammatical or punctuation (dot, comma, capitalization) errors.	Correct choice of vocabulary **(the ones taught for this writing task)** and no spelling errors. The register or style consistently matches the nature of the task.
4	Sufficiently addresses all parts of the task. Presents a developed response with supported ideas.	Arranges information and ideas coherently and there is overall progression. **The topic sentence and the supporting sentences are well-developed.** Uses discourse markers listed in the syllabus (and, but, so, because) though there may be some incorrect use. Paragraphing may not be logical.	Uses most of the structures **required for the task (present simple, present continuous, adverbs of frequency)** with only rare grammatical or punctuation (dot, comma, capitalization) errors which do not impede communication.	Only rare spelling errors or incorrect choice of vocabulary **(the ones taught for this writing task)** which do not impede communication. The register or style may in occasional instances not match the nature of the task.
3	Addresses most parts of the task. Presents ideas which may not be fully developed and/or lack focus.	Presents information with some organization but there may be a lack of overall progression. **The topic sentence and supporting sentences are partially developed but do not represent a well-developed paragraph.** Uses only a limited number of discourse markers listed in the syllabus (and, but, so, because) and there may be frequent incorrect use. Does not divide ideas into individual paragraphs.	Uses only some of the **structures required for the task (present simple, present continuous, adverbs of frequency)** with some grammatical or punctuation (dot, comma, capitalization) errors which may impede communication.	Some spelling errors and incorrect choice of vocabulary **(the ones taught for this writing task)** which may impede communication. The register or style frequently does not match the nature of the task.

Mark	TASK RESPONSE	ORGANIZATION	GRAMMAR	VOCABULARY
2	Addresses some parts of the task. Presents ideas which may be inadequately developed.	Ideas are not arranged coherently and there is no clear progression in the response. **The topic sentence or/and supporting sentences are not coherently developed.** Uses very few discourse markers listed in the syllabus (and, but, so, because) and their use is usually inaccurate or repetitive.	Uses very few of the structures required for the task **(present simple, present continuous, adverbs of frequency)** with frequent grammatical or punctuation (dot, comma, capitalization) errors which usually impede communication because of lack of grammatical elements.	Frequent spelling errors and incorrect choice of vocabulary **(the ones taught for this writing task)** which usually impede communication. The register or style usually does not match the nature of the task.
1	Addresses the task only minimally.	Has very little control of organizational features **to make the topic sentence and develop the related supporting ideas.** Makes no attempt to use discourse markers listed in the syllabus (and, but, so, because).	The range of structures (present simple, **present continuous, adverbs of frequency)** used is extremely limited and grammatical and punctuation (dot, comma, capitalization) errors are so prevalent that hardly any communication takes place.	Spelling errors and incorrect choice of vocabulary **(the ones taught for this writing task)** occur to such an extent that hardly any communication takes place. The register or style does not match the nature of the task at all.

1. Give a zero if the task is not attempted, if the answer is totally incomprehensible or if the answer is plagiarized and there is evidence for this.
2. If the answer is totally irrelevant and unrelated to the task in any way, award a zero for the Task Response, and don't give more than 2 marks for Organization, 2 for Grammar and 1 for Vocabulary.
3. Candidates are penalized for writing fewer than the required number of words. The reduction is from Task Response as follows:
    - 60 words = -3;
    - 61–100 words = -2;
    - 101–140 = -1
4. If the word count is less than 50% do not award more than 3 marks for the other three criteria.
5. To receive the mark allocated for a criterion, all positive features mentioned in the descriptors should be achieved.

- Task Response: The writing is not off topic and covers all areas of the question content.

Authors' Biography

Zahra Zargaran holds a Ph.D. in TEFL from Azad University, Science and Research Branch, Tehran, Iran. She is a lecturer at University of Technology and Applied Sciences, Shinas, Oman and is also the coordinator of Academic Advising Committee. She is a certified trainer by the British Council, IDP Australia, and Cambridge and has actively trained hundreds of teachers in both public and private sectors nationally and internationally since 2011. With 19 years of experience in teaching and teacher training, Dr. Zargaran specializes in language education, curriculum design, materials development, and professional skills training. She has conducted many research training workshops and courses and is an active researcher in the areas of language teaching/learning, educational assessment, teacher education, learning strategies, teacher cognition, AI in education and IELTS teaching.

Mohsen Ghorbanpoor received his B.A. and M.A. degrees in English Language Teaching from Kharazmi University and Tarbiat Modares University, Tehran, respectively. He began his teaching career in 2010 and has taught English in both public education and private institutes in Iran. In 2020, he joined the University of Technology and Applied Sciences in Al Mussanah, Oman, where he currently serves as an ELT lecturer and the coordinator for professional development and research. His areas of interest include teacher education, materials development, listening, writing and digital pedagogy. He has also conducted in-house training programs and coordinated institutional research activities.

Volume 14, Issue 1
Winter and Spring, 2026
pp. 31-52

Journal of Applied Linguistics and Applied Literature: Dynamics and Advances

Form-Focused Tasks in Task-Based Language Teaching for Enhancing EFL Learners' Speaking Fluency and Motivation

Simin Anbarshahi [1, *] and Omid Yeganeh [2]

[1] *Corresponding author: Assistant Professor of TEFL, English Department, Human Sciences, Azarbaijan Shahid Madani University, Tabriz, Iran*
Email: anbarshahi@gmail.com

[2] *MA in ELT, English department, Azarbaijan Shahid Madani University, Tabriz, Iran, Email: omidyeganeh44@yahoo.com*

Abstract

Task-Based Language Teaching (TBLT), as a prominent pedagogical approach, has emerged as a significant innovation in language education, particularly recognized for its potential to enhance speaking fluency among Iranian EFL learners and aims to enhance accuracy, fluency and complexity of the utterances produced by language learners. The purpose of the present study is to examine the impact of TBLT on learners' motivation and speaking fluency. The research design included both experimental and control groups, each participating in a 12-session course. The experimental group engaged in TBLT with an emphasis on form, while the conventional methods were utilized for the control group. Speaking fluency was evaluated using Rate A and Rate B fluency measures and academic motivation was assessed through a motivation questionnaire based on Dörnyei and Dewaele's (2002) L2 Motivational Self System Questionnaire, which included two main sections. A speaking assessment was administrated as a pretest before the treatment and as a posttest after the treatment. Afterward, the result was analyzed; the findings indicated that the experimental group, which received TBLT, outperformed the control group in both motivation and speaking fluency. The results suggest that TBLT, particularly through real-world tasks, encourages students to practice speaking in meaningful contexts, making the learning process more engaging. As a result, students exhibit increased motivation and readiness for lessons, highlighting the potential of TBLT to enhance both language proficiency and learner motivation. This study offers significant insights for language educators, course designers, and policymakers in the field of EFL, suggesting that task-based activities should be incorporated into language learning curricula to improve not only linguistic proficiency but also student motivation and engagement.

Keywords: speaking fluency, task-based language teaching, motivation, Focus on Form

ARTICLE INFO

Research Article
Received: Wedensday, May 28, 2025
Accepted: Thursday, March 12, 2026
Published: Wedensday, April 1, 2026
Available Online: THursday, March 12, 2026
DOI: https://doi.org/10.22049/jalda.2026.30662.1813

Online ISSN: 2821-0204; Print ISSN: 28208986

Introduction

The ability to speak effectively is essential in language acquisition, enabling students to engage in natural and spontaneous conversations with native speakers. Thoughtfully crafted speaking activities in English as a Foreign Language (EFL) setting can significantly boost students' motivation and create a vibrant and interactive learning atmosphere (Celce-Murcia, 2001; Nunan, 1999; Ochoa-Cueva et al., 2016). Nevertheless, both EFL educators and students encounter considerable obstacles in enhancing speaking skills (James et al., 2019). Various factors contribute to these difficulties. EFL learners frequently find themselves in environments where English is seldom used, which restricts their chances to practice outside the classroom. Furthermore, many students depend predominantly on coursebooks as their main or only resource (Akbari, 2015), hindering their ability to transfer classroom knowledge to real-world situations. This reliance can result in diminished motivation and the belief that speaking English holds little practical significance (Littlewood, 1984). Additionally, as highlighted by Al-Hosni (2014), speaking skills often do not receive adequate focus in language programs, particularly in educational institutions, resulting in fluency challenges for both educators and students, especially within public education systems. Lastly, teacher-centered classroom dynamics, characterized by excessive Teacher Talk Time (TTT), diminish students' chances for active engagement and speaking practice.

The strategies typically employed to enhance EFL learners' speaking abilities often overlook direct speaking practice, which hampers learners' effective acquisition and application of this skill (Al-Hosni, 2014). This challenge is also prevalent among Iranian EFL learners. Despite the introduction of various contemporary teaching methods in Iran's educational framework—such as discussions, role-plays, simulations, information gap activities, brainstorming, interviews, reporting, and card games (Staab, 1992)—these approaches often fail to cultivate speaking proficiency. Afshar and Asakereh (2016) identify a significant drawback of traditional language learning methods: they focus on memorized phrases rather than genuine language application, which hinders the improvement of communicative competence.

Another crucial element that aids EFL learners in language acquisition and improving the communication abilities is motivation. Keller (1983) characterizes motivation as the decisions individuals make about which experiences or objectives to pursue or avoid, along with the effort they invest in achieving those objectives. Numerous studies have demonstrated that motivation is crucial for successful language acquisition, irrespective of the students' gender, age, or proficiency level (Alrabai & Moskovsky 2016; Dörnyei & Ottó, 1998). The domain of language education underscores the importance of learner motivation in both mastering the target language and engaging with its cultural context (Henry & Thorsen, 2018). Lam and Law (2007) assert that students are more motivated when educators assign challenging tasks, integrate real-life activities, stimulate curiosity for new knowledge, provide opportunities for independence, recognize students' efforts, and deliver constructive feedback. Recent studies similarly highlight that motivation is strengthened when students envision themselves as competent future L2 speakers

(Al-Hoorie & Dörnyei, 2017). TBLT is acknowledged as an approach that boosts students' motivation to communicate fluently in English by prioritizing language use within the context of task completion. TBLT encourages authentic communication, thereby enhancing the learning process (Ahmadian & Long, 2020). While completing tasks, students participate in meaning negotiation through discussions. Furthermore, authentic, real-world tasks—an essential aspect of communicative activities—actively engage students and inspire them to utilize the target language (Ellis, 2003). These interactive teaching strategies cultivate a social and engaging classroom atmosphere that promotes meaningful language practice.

Ellis (2003) characterized TBLT as a dynamic extension of the CLT methodology, which prioritizes meaningful, form-focused, and real-world tasks. This approach underscores the active participation of students, connecting their involvement to their personal attitudes and levels of motivation. Ballesteros and van Compernolle (2023) posited that task planning within TBLT framework has important implications for speaking development. Dörnyei (1998) pointed out that both educators and researchers consider motivation to be a vital element affecting the speed and efficacy of second or foreign language (L2) acquisition. As a result, there exists a considerable amount of research investigating the correlation between TBLT and learner motivation in countries like Spain and Singapore. In Iran, TBLT as a practical and effective pedagogical approach, offers learners valuable opportunities to engage in meaningful and real-world tasks that foster communication and enhance motivation (Mahdavirad, 2017). Tale and Goodarzi (2015) noted that TBLT enables learners to acquire knowledge through active participation in tasks. The interplay between motivation and speaking fluency has long been recognized as reciprocal and dynamic in second language acquisition.

Motivation not only drives learners to engage more actively in communicative tasks but also sustains their effort in developing fluent speech (Dörnyei, 2005; Ushioda, 2011). Conversely, improvements in speaking fluency can enhance learners’ self-confidence and intrinsic motivation, creating a positive feedback loop that accelerates language development (MacIntyre et al., 2011). Within the TBLT framework, where authentic communication and meaningful interaction are central, motivation may directly influence the extent to which learners participate in tasks, while fluent task performance may, in turn, reinforce their motivation. Despite this theoretically grounded relationship, empirical research exploring the correlation between speaking fluency and motivation among Iranian EFL learners within TBLT contexts remains scarce. Therefore, examining this link can offer valuable insights into how affective and linguistic factors interact in language learning.

Over the years, TBLT has gained traction in language education and has drawn considerable scholarly interest (Ellis, 2003; Skehan, 1996; Nunan, 2004). In this context, the enhancement of speaking fluency among Iranian EFL learners has recently become a prominent topic of discussion. While numerous studies have established a positive link between motivation and various facets of foreign language learning, the specific impact of TBLT on EFL learners' motivation to enhance their speaking abilities remain underexplored.

In light of the shifting perspective on language learning as both a psychological and technological process, this study seeks to explore the effects of Task-Based Language Teaching (TBLT), particularly with an emphasis on form, on two critical outcomes: motivation and speaking fluency. It will specifically assess whether TBLT, when focusing on form, significantly influences speaking fluency and motivation of Iranian EFL learners. Its significance lies in its focus on the Iranian context, providing valuable insights into the possible benefits of this instructional approach. Additionally, the study intends to fill a gap in the literature and offer practical recommendations for EFL educators in Iran to improve teaching practices and learning outcomes. In light of the aforementioned points, the research questions have been proposed as follows:

1. Does TBLT approach focusing on form affect on Iranian EFL learners' speaking fluency significantly?
2. Does TBLT approach focusing on form affect on Iranian EFL learners' motivation significantly??
3. Is there any correlation between learner' speaking fluency and motivation?

Based on the above questions, the following hypotheses were formed:

1. TBLT approach focusing on form affect on Iranian EFL learners' speaking fluency significantly.
2. TBLT approach focusing on form affect on Iranian EFL leaners' motivation significantly.
3. There is a correlation between learners' speaking fluency and motivation.

Literature Review

Theoretical Framework

This study is grounded in the principles of Task-Based Language Teaching (TBLT), Focus on Form (FonF), and motivation in second language acquisition. TBLT emphasizes the use of authentic, real-world tasks to promote meaningful communication, enabling learners to produce language in contexts that mirror natural interactions (Hismanoglu & Hismanoglu, 2011; Richards & Rodgers, 2001). By integrating form-focused attention within communicative tasks, TBLT supports both fluency and linguistic accuracy, ensuring learners develop functional proficiency alongside grammatical competence (Ellis, 2003; Mystkowska, 2011).

Motivation plays a critical role in language learning, as intrinsically motivated learners engage more actively, persist longer, and practice language more effectively (Masrom et al., 2015). TBLT inherently fosters motivation by involving learners in collaborative, meaningful tasks that stimulate creativity, autonomy, and sustained engagement (Nita et al., 2019; Willis & Willis, 2007). Accordingly, this study adopts a task-based, form-focused, and motivation-oriented framework, positing that well-designed tasks enhance both speaking fluency and learner motivation, providing a theoretical rationale for examining the impact of TBLT on Iranian EFL learners.

Motivation

Motivation is essential in language acquisition, and TBLT has been recognized as an effective method for enhancing learners' motivation to study and utilize English. TBLT inherently involves students in meaningful and authentic language activities that foster active engagement and creativity. Research conducted by Masrom et al. (2015) indicated that motivated learners participating in these tasks were able to produce language that was more lexically complex compared to their less motivated counterparts. These activities encourage creativity and interaction, especially through collaborative pair and group work, where learners apply language in genuine, spontaneous contexts. Consequently, TBLT cultivates a learning atmosphere that significantly enhances motivation within EFL classroom.

Recent research underscores the crucial role of meaningful, well-structured tasks in fostering learners' motivation and engagement in language learning. Scholars such as Dörnyei and Ushioda (2011) and Ellis (2017) emphasize that tasks designed to promote autonomy, relevance, and interaction can enhance learners' intrinsic motivation, which in turn supports their sustained effort and achievement. This highlights the pedagogical importance of incorporating motivating, communicative tasks into English instruction. Additionally, offering an authentic social context for using English in real-life scenarios is crucial, as it not only boosts motivation but also aids learners in reaching their educational objectives and enhancing their language skills (Kumaravadivelu, 2006).

Task- Based Language Teaching

TBLT has emerged as a powerful instructional strategy aimed at improving language learners' communication abilities, particularly in speaking. In contrast to conventional methods that primarily emphasize on grammar, TBLT prioritizes real-life tasks, allowing learners to engage in authentic interactions. This approach encourages students to articulate their thoughts, confirm understanding, seek clarification, and adapt to the language they encounter, which may surpass their current proficiency level. Hismanoglu and Hismanoglu (2011, p. 48) observe that the discourse generated from tasks is intended to closely resemble the natural communication that occurs in real-life situations.

TBLT was created to overcome the shortcomings of traditional language teaching methods, which frequently emphasize grammar at the expense of communication. This grammar-centric approach can hinder learners from attaining a high degree of proficiency and fluency in the target language. Richards and Rodgers (2001), argue that conventional methods have inadequately prioritized speaking, despite its significance as the most direct and accessible means of communication. Even among motivated and capable students, those instructed through traditional structural methods often find it challenging to achieve optimal proficiency and fluency in a second language, even after years of study (Skehan, 1996). Notably, TBLT does not dismiss the traditional emphasis on form; rather, it combines both communicative and form-focused aspects, providing a well-rounded approach to language acquisition (Ellis, 2003). Shirazifard et al. (2023) argued that TBLT-synthesized collaborative dialogue had significant effect on improvements in language production.

Focus on Form

Speaking fluency is frequently regarded as a crucial measure of success for students learning a foreign language, as the main objective of communication is to ensure that the listener comprehends the speaker's intended message. Ellis (2005) characterizes fluency in task performance as the degree to which the language produced during a task exhibits pausing, hesitation, or reformulation. However, in the past two decades, educators have debated whether the emphasis should be placed on accuracy and grammatical structure or on fluency and meaning in language learning. The issue of how much focus should be given to grammar in communicative classrooms remains contentious (Doughty & Williams, 1998; Lightbown, 2000; Norris & Ortega, 2000). Much of the discussion revolves around the extent to which teachers should concentrate on grammatical forms during instruction. Mystkowska, 2011) delved deeper into the function of incidental Focus on Form (FonF) and explored the relationship between classroom participation structures and the quantity, nature, and effectiveness of FonF. Their research found a clear correlation between the structure of classroom participation and the success of FonF, highlighting the significance of integrating form-focused instruction in communicative contexts.

While FonF is traditionally used to teach grammar, research shows it can also support speaking fluency when integrated into meaningful communicative tasks. In TBLT, form-focused attention helps learners notice gaps in their language production and refine their output without disrupting communication, reducing hesitations and reformulations (Ellis, 2003; Mystkowska, 2011). Research indicates that integrating FonF within meaningful tasks enables learners to improve linguistic accuracy without undermining communicative objectives (Salimi et al., 2015), thereby supporting both fluency and correctness in speaking (Ellis, 2005). Thus, in our study, FonF was applied as a scaffolding tool within tasks to enhance fluent, confident, and continuous language production, aligning with evidence that FonF can simultaneously support accuracy and fluency.

Experimental Studies

Over the last four decades, a wide range of techniques and methodologies have emerged to support learners in enhancing their speaking skills (Derakhshan et al., 2015). The transition to communicative approaches in the 1980s brought about significant changes in attitudes and methods of language teaching, leading to the creation of strategies that prioritize speaking proficiency, such as TBLT.

Speaking Fluency and TBLT

Recent experimental and action-research studies provide strong evidence that TBLT can enhance learners' speaking fluency. For example, in a study of 20 EFL learners at a private university in Colombia, ten fluency-based tasks were implemented and qualitative and quantitative data indicated reductions in pauses, improved speech rate, and smoother pacing. Learners reported positive perceptions of TBLT as a methodology that supported their spoken expression and confidence (Zúñiga et al., 2023). Similarly, in an action-research project involving 25 in-service teachers of English in Ecuador (A2 level), post-test results showed statistically significant improvements in oral interaction after task-based activities: mean scores rose from 17.72 to 19.12 ($p < .001$). Qualitative data indicated that participants felt their confidence increased and anxiety decreased during speaking (Detken, et al., 2024). Another large-scale study involving 60 intermediate-level ESL learners over six weeks found that the TBLT group showed significantly greater enhancement in speech rate, lexical diversity, and reductions in hesitation compared to a conventional grammar-based group. Qualitative findings noted higher levels of motivation and confidence among the TBLT group (Kholbutayeva, 2025).

Additionally, Nget et al. (2020) investigated the impact of task-based instruction on the speaking skills of 9th-grade students. The data gathered from speaking assessments, both quantitative and qualitative, demonstrated that task-based teaching significantly improved students' speaking skills by creating a more supportive learning environment, enhancing their confidence, and increasing their motivation. In a similar vein, Hashemifardnia et al. (2019) conducted research involving 50 Iranian EFL students, who were split into experimental and control groups. The findings indicated that task-based activities markedly improved the speaking fluency of the experimental group, particularly as the tasks reflected real-life contexts, enabling students to practice and refine their fluency.

These studies collectively indicate that task-based instruction — particularly when tasks simulate real communicative demands — can foster greater fluency in speaking, presumably by increasing meaningful output, reducing planning time, and providing more opportunities for spontaneous interaction.

Motivation and TBLT

Furthermore, motivation, a vital element in language acquisition, has been positively impacted by TBLT. By involving students in meaningful tasks that simulate real-life scenarios, TBLT cultivates an interactive and motivating educational atmosphere. This inherent motivation is further enhanced as students interpret meaning and engage in communication while completing tasks (Willis & Willis, 2007). Supporting these findings, Nita et al. (2019) investigated the effect of TBLT on the speaking abilities of EFL learners who possess intrinsic motivation. The research, carried out at a public senior high school in Padang, compared an experimental group taught through TBLT with a control group instructed using traditional methods. The findings indicated that students in the experimental group, who were intrinsically motivated, exhibited greater speaking proficiency than those in the control group, underscoring the effectiveness of TBLT in improving both speaking skills and motivation.

Further evidence comes from studies that report improved learner engagement, willingness to communicate, and reduced anxiety when tasks are meaningful and communicative rather than purely accuracy-focused. For instance, surveys in recent research indicate that learners perceive TBLT as more enjoyable, challenging in a positive way, and confidence enhancing — all of which point to elevated motivation. (Nugrahaeni, 2022). Thus, the literature suggests that well-designed tasks in TBLT not only promote speaking fluency but also contribute to an interactive, motivating classroom climate — an outcome especially relevant for your study on the dual focus of fluency and motivation.

Focus Within TBLT: Gap & Justification

While the general effectiveness of TBLT for speaking fluency and motivation is affirmed, fewer studies explicitly focus on form-focused tasks embedded within the TBLT framework (i.e., tasks that integrate attention to form within communicative tasks). Some research mentions form-focus implicitly (e.g., during the language-focus phase of the task cycle) but does not systematically examine its effect on fluency and motivation together. For example, one recent study pointed out that tasks enabled learners to focus on meaning and in doing so reduced hesitations, but did not isolate the effect of form-focused intervention (Zhang, 2025)

Consequently, by explicitly combining form focused instruction within authentic tasks, this research advances understanding of how accuracy-oriented attention (form) and communicative fluency (meaning) interact to influence speaking proficiency and motivation.

In sum, the body of experimental research supports TBLT as an effective pedagogical approach for enhancing speaking fluency and learner motivation.

Method

Participants

This study utilized a quasi-experimental design. The original population consisted of approximately 120 adolescent EFL learners enrolled at the private language institute in Tabriz, East Azerbaijan, Iran. These learners attended general English courses at various proficiency levels, ranging from elementary to upper-intermediate. From this population, 42 students in two intact classes at the intermediate level were selected for the study and were randomly assigned to the experimental and control groups. The participants included both male and female learners aged 13 to 18 years. To ensure homogeneity in English proficiency, the Oxford Placement Test (OPT) was administered, and 34 students whose scores fell within the intermediate range were retained for analysis. Furthermore, a speaking assessment was conducted to confirm that all participants had a comparable level of speaking ability and met the study's criteria. Subsequently, the participants were randomly divided into two groups as control group and experimental group. The experimental group underwent treatment aimed at enhancing speaking fluency. Both groups completed pretest and posttest speaking assessments to measure the effect of the intervention. This treatment utilized TBLT with an emphasis on form.

Materials and Instruments

The initial instrument utilized in this research was the OPT, which facilitated the evaluation of participants' skill levels, allowing for the selection of individuals who met the intermediate proficiency standard while excluding those who did not qualify. This assessment is specifically designed to measure proficiency in a S/F language, focusing on both grammatical and pragmatic knowledge, as well as the capacity of effectively utilize this knowledge in communication. Its validity and reliability are well-documented, making it a widely accepted tool for categorizing students into different proficiency levels based on their performance (Weiss, 2011). The second tool employed was a speaking assessment developed by the researcher, tailored to the materials the students were studying during that semester. This test functioned as a pretest to identify any significant differences in the learners' speaking abilities before the intervention. It was also administered as a post-test to evaluate any statistically significant enhancements in the oral performance of the experimental group following the treatment. Both pre and posttests had the same topics and there weren't any differences.

The motivation questionnaire comprised items pertaining to the participants' motivation, structured according to the framework proposed by Dörnyei (2002). It was administrated in English version and was organized into two primary sections. the first section consists of items measuring the learners' attitudes and motivation concerning English learning, and the second section consists of questions about the learners' background information. The questionnaire included closed-ended items and they were accompanied by six response options for respondents to indicate the extent to which they agree or disagree with it by marking one of the responses referred as Likert scale. The pretest and posttest of speaking fluency were rated by two raters. Pearson correlations were computed to probe the inter-rater reliability indices of the two raters. The results indicated that the values were higher than .50, it was concluded that the inter-rater reliability indices enjoyed large effect sizes.

The internal consistency or reliability of the questionnaire is measured by the reliability coefficient. The reliability indices for pretest and posttest of motivation were .941, and .953 respectively. Based on the criteria, it can be concluded that the reliability indices for pretest and posttest of motivation were "excellent".

Furthermore, the study included listening files from the students' textbook at the institute, which provided the basis for the speaking practice materials. These resources were utilized to assess and improve the students' speaking fluency.

Procedure

The experimental group underwent a procedure divided into three distinct phases:

Step 1: Preview (Pre-Task Phase) – This initial phase concentrated on form, aiming to introduce the subject matter while activating the students' existing knowledge.

Step 2: Listening and Speaking (Task Phase) – In this phase, the focus remained on form, with the goal of completing the task and using the target language structures in context. This step is sometimes addressed to as the task cycle phase.

Step 3: Acting out or Role-Playing – In this final phase, the focus was again on form, to consolidate and review the target language structures. This phase is often addressed to as the language focus phase. During this stage, students engaged in pair work to practice dialogues, simulating real-life scenarios.

In contrast, the control group adhered to a more conventional method, where the teacher presented the topics, followed by students practicing dialogues as a production activity, akin to rote learning, such as memorizing scripts, emphasizing explicit grammar explanation, and memorization of model dialogues from the textbook. Students practiced predetermined patterns in limited, accuracy-focused exercises, with minimal opportunities for spontaneous or communicative use of the target language. This teacher-centered method contrasted with the task-based, form-focused instruction in the experimental group, which emphasized interaction and meaningful communication through real-life tasks. After a six-week instructional period, both groups participated in a speaking post-test to evaluate the effects of TBI on their speaking fluency. Furthermore, a questionnaire of motivation was re-administered to assess any changes in the students' motivation levels resulting from the TBLT approach.

In this study, the instructional treatment followed the principles of Task-Based Language Teaching (TBLT), not Task-Supported Language Teaching (TSLT). The tasks designed for the experimental group were primarily open-ended authentic, real-life communicative tasks that required learners to use English meaningfully to achieve specific outcomes, such as information exchange, role-play, and problem-solving. Tasks were carefully sequenced from simpler to more complex communicative demands. In the pre-task phase, learners were introduced to the topic and relevant language forms, activating prior knowledge. During the task cycle, students engaged in paired and group activities requiring them to use the target language spontaneously. Finally, in the language focus phase, students consolidated linguistic structures through guided practice and peer interaction, reinforcing fluency while maintaining attention to form. While the tasks were not graded in a conventional sense, progression and complexity were monitored across the six sessions. Criteria for sequencing included linguistic complexity, cognitive demand, and communicative challenge, ensuring learners gradually built confidence and fluency.

While the primary emphasis was on meaning and fluency, a Focus on Form (FonF) was incorporated implicitly and reactively during task performance—drawing learners' attention to relevant linguistic forms that naturally emerged from communication. This integration of form-focused attention within real-world tasks allowed learners to enhance both speaking fluency and motivation while maintaining the communicative purpose of the activity.

Data Analysis

Fluency was analyzed using documented procedures widely employed in second language research (e.g., Ellis & Barkhuizen, 2005; Skehan, 2009; Tavakoli & Skehan, 2005;). These metrics encompassed speech rate, pauses, word frequency, vocabulary richness, and grammatical complexity. To assess fluency, two key factors were considered:

- Rate A: The number of syllables produced per minute of speech
- Rate B: The number of meaningful syllables produced per minute of speech

These metrics were selected based on established procedures for measuring L2 oral fluency, ensuring that the analysis followed validated and reliable standards.

For the motivation data, the questionnaire scores were analyzed using Cronbach's alpha in the SPSS software. A value greater than 0.70 was considered acceptable to ensure reliability of the responses. Following this, the hypotheses were tested using One-Way ANCOVA on the post-test scores from both groups to evaluate the treatment's effect. A p-value of less than 0.05 was required to establish a significant difference between the results of the two groups.

Inter-Rater Reliability Indices

To probe the inter-rater reliability related to the two raters for the pretest and posttest of speaking fluency, Pearson correlations were calculated. The findings show a large agreement between the two raters for both the pretest (r (30) = .662, representing a significant effect, $p < .05$) and the posttest (r (30) = .815, representing a large effect size, $p < .05$) of speaking fluency. These findings suggest that both the pretest and posttest had strong inter-rater reliability.

Cronbach's Alpha Reliability Indices: The motivation questionnaire demonstrated high internal consistency for both pretest and posttest administrations. Cronbach's alpha coefficients were .941 for the pretest and .953 for the posttest, indicating satisfactory reliability of the instrument.

Results

In this research, ANCOVA was used for analyzing the effects of TBLT on the dependent variable, while controlling for the influence of other variables. To ensure the validity of the analysis, normality was assessed through skewness and kurtosis for the pretest and posttest of speaking fluency and motivation. The skewness and kurtosis values for the tests confirm the normal distribution and reliability of the data.

First, regarding the One-Way ANCOVA, it is important to note that a linear correlation between the pretest and posttest scores of speaking fluency is required. The results of the linearity test (F (1, 20) = 31.67, $p < .05$, eta squared = .686, representing a significant effect size), showed a significant linear correlation between the pretest and posttest of speaking fluency, thus rejecting the null hypothesis concerning the correlation between these two tests was non-linear.

Second, the assumption related to homogeneity of regression slopes was tested for One-Way ANCOVA, which requires the consistency of the linear relationship between the pre and post-tests of speaking fluency across the two groups. After examining the assumption, t was shown that the interrelationship between the covariate (pretest) and the independent variable was found to be non-significant ($F(1, 28) = .031$, $p > .05$, Partial eta squared = .001, representing a weak effect). Therefore, the null hypothesis, suggesting a non-linear relationship between the pretest and posttest of speaking fluency of the two groups was rejected by indication of this result.

Finally, after accounting the pretest effect, One-Way ANCOVA assumes that the variances on the posttest of the two groups' speaking fluency are approximately equal. This assumption is referred as the homogeneity of variances. The non-significant results obtained from Levene's test ($F(1, 30) = 1.28$, $p > .05$), indicated that this assumption was upheld.

After reviewing the assumptions for One-Way ANCOVA, the key findings include descriptive statistics and the primary outcomes of the analysis. Table 1 presents the mean scores of both groups on the posttest of speaking fluency, controlled for any influence of pretest. The findings indicate that the experimental group ($M = 46.96$, $SE = 2.95$) outperformed the control group ($M = 37.53$, $SE = 2.95$) on the posttest of speaking fluency, adjusted for the pretest scores.

Table 1

Descriptive Statistics of Speaking Fluency Posttest *by Group with Pretest*

Group	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Experimental	46.966[a]	2.955	40.922	53.009
Control	37.534[a]	2.955	31.491	43.578

a. The values of Covariates in the model are as follows: Pretest = 27.59.

Table 2 presents the results of the One-Way ANCOVA. The findings ($F(1, 29) = 36.89$, $p < .05$, partial $\eta^2 = .148$, reflecting a strong effect) reveal that the experimental group performed at higher level than the control group on the post-test of speaking fluency after accounting the pre-test scores. Consequently, the null hypothesis supposing that "TBLT focus on forms did not significantly improve fluency in EFL learners' speaking ability" was rejected.

Table 2

Between-Subjects Effects Tests for Posttest of Speaking Fluency by Groups with Pretest

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Pretest	5095.091	1	5095.091	36.899	.000	.560
Group	695.334	1	695.334	5.036	.033	.148
Error	4004.378	29	138.082			
Total	66469.000	32				

To assess the second null hypothesis, the assumption underlying the One-Way ANCOVA were systematically evaluated. The linearity test yielded a statistically significant results (F (1, 18) = 77.08, p < .05, eta squared = .819, presented in Table 3) indicated a large effect size, leading to the rejection of the null hypothesis, and confirm that there was linear relationship between the pretest and posttest of motivation

Table 3

Linearity Between Pretest and Posttest of Motivation by ANOVA Test

			Sum of Squares	df	Mean Square	F	Sig.
PostMotivation* PreMotivation	Between Groups	(Combined)	44.827	13	3.448	6.247	.000
		Linearity	42.550	1	42.550	77.089	.000
		Deviation from Linearity	2.277	12	.190	.344	.968
	Within Groups		9.935	18	.552		
	Total		54.762	31			
	Eta Squared		.819				

The interaction effect between the covariate (pretest) and the independent variable was not statistically significant (F (1, 28) = 3.64, p > .05, the observed Partial eta squared value of .115, indicate an effect of intermediate size, leading to the rejection of the null hypothesis. This outcome confirme that the relationship between the pretest and posttest of motivation was indeed linear across both groups.

Table 4

Testing Homogeneity of Regression Slopes for Motivation Posttest by Groups with Pretest

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Group	.294	1	.294	2.889	.100	.094
Pretest	42.550	1	42.550	417.570	.000	.937
Group* Pretest	.371	1	.371	3.645	.067	.115
Error	2.853	28	.102			
Total	307.579	32				

Given the non-significant findings from Levene's test (F (1, 30) = .371, p > .05), the assumption of homogeneity of variances across groups was upheld (Table 5).

Table 5

Levene's Test of Homogeneity of Variances for Motivation Posttest by Groups with Pretest

F	df1	df2	Sig.
.371	1	30	.547

Table 6 presents the means of adjusted posttest motivation for the experimental and control groups, accounting for the effect of the pretest. The findings revealed that the experimental group (M = 3.34, SE = .083) demonstrated a higher adjusted mean on the motivation post-test compared to the control group (M = 2.28, SE = .083), after accounting for the pretest differences.

Table 6

Descriptive Statistics for Posttest of Motivation by Group with Pretest

Group	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Experimental	3.341[a]	.083	3.170	3.511
Control	2.281[a]	.083	2.110	2.451

a. Covariates appearing in the model are evaluated at the following values: Pretest = 1.62.

And finally, table 7 indicates the main outcomes of One-Way ANCOVA. The findings ($F(1, 29) = 80.82$, $p < .05$, partial $\eta^2 = .736$, indicating a higher effect) approve the significant outperformance of experimental group over the control group on the motivation posttest, after adjusting for the pretest, leading to the rejection of the second null hypothesis, which stated that "TBLT focus on forms did not significantly improve fluency in EFL learners' speaking ability."

Table 7

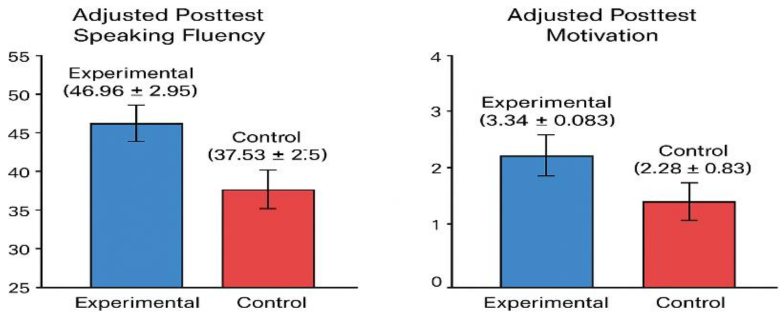
Tests of Between-Subjects Effects for Motivation Posttest by Groups with Pretest

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Pretest	42.550	1	42.550	382.670	.000	.930
Group	8.988	1	8.988	80.829	.000	.736
Error	3.225	29	.111			
Total	307.579	32				

Based on the result obtained from the data, adjusted Posttest Means of Speaking Fluency and Motivation by Groups were presented in Figure 1.

Figure 1

Adjusted Posttest Means by Group



As shown in table 8, the findings of the Pearson Correlation (r (30) = .878, representing a large effect size, p < .05) revealed a strong correlation between motivation and speaking fluency. Therefore, the third null hypothesis, stating "there is no significant relationship between motivation of EFL learners and their speaking fluency," was rejected.

Table 8

Pearson Correlation between Motivation and Speaking Fluency

		Speaking Fluency
Motivation	Pearson Correlation	.878**
	Sig. (2-tailed)	.000
	N	32

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

Discussion

This study aimed to examine the effects of TBLT emphasizing the form on the speaking motivation and fluency of Iranian EFL learners. It specifically investigated whether this teaching method significantly influenced speaking fluency and motivation of learners, as well as the potential correlation between these two factors. By delving into these elements, the research sought to enhance understanding of TBLT's effectiveness in advancing language proficiency, while also taking into account the psychological and motivational aspects that are vital in the process of language learning. The findings revealed a notable improvement in the speaking fluency of the experimental group in the comparison with the control group. The experimental group, which participated in TBLT with a focus on form, exhibited more substantial advancements in their speaking skills. This result supports the initial hypothesis that “The TBLT approach focusing on form has a significant effect on Iranian EFL learners' speaking fluency." The observed enhancements in speaking fluency, attributed to TBLT's focus on task completion and form-oriented practice, highlight the approach's effectiveness in fostering speaking proficiency among Iranian EFL learners.

In terms of motivation, the results also confirmed the second hypothesis: learners exposed to TBLT showed a noticeable increase in motivation to learn and use English. The dynamic and engaging nature of task-based activities appears to have contributed to this heightened motivation. Through active participation in meaningful, goal-oriented tasks, students developed a sense of ownership over their learning. The interactive group work, peer feedback, and focus on real-life communication likely created a more stimulating environment that made learning English more enjoyable and relevant. These findings highlight TBLT's capacity to enhance intrinsic motivation by fostering autonomy, competence, and relatedness—three key psychological needs identified in motivational theory.

Moreover, a strong positive correlation was found between speaking fluency and motivation, supporting the third hypothesis. This relationship indicates that as students' fluency improves, they become more motivated to continue learning, and conversely, higher motivation encourages more engagement in communicative activities, which in turn enhances fluency (Naseri et al., 2025). This reciprocal effect underscores the importance of addressing both cognitive and affective dimensions in language pedagogy, as proficiency and motivation appear to reinforce one another in a cyclical process of growth.

Overall, the findings suggest that TBLT with a focus on form effectively promotes both speaking proficiency and motivation among Iranian EFL learners. The observed improvement may stem from the interactive, learner-centered, and collaborative nature of TBLT tasks, which foster meaningful communication and peer learning. By prioritizing communicative tasks over isolated grammar drills, learners internalize language more deeply and retain it more effectively. The inclusion of real-life, contextually relevant tasks also helped connect classroom learning to students' everyday experiences, sparking greater interest and engagement. Furthermore, the clear communicative purposes of the tasks, combined with positive reinforcement from teachers and peers, further enhanced learners' confidence and motivation.

The improvements can be attributed to the interactive and authentic nature of TBLT tasks. Role-playing, communicative activities, and tasks reflecting students' daily experiences encouraged meaningful practice, peer collaboration, and active participation. These tasks emphasized fluency over grammatical perfection, allowing students to communicate freely, build confidence, and internalize language structures more effectively. Positive feedback from teachers and peers further enhanced motivation and engagement. These findings are consistent with previous studies highlighting TBLT's effectiveness in enhancing communicative competence, learner engagement, and a supportive learning environment (Nget et al., 2020; Willis & Willis, 2007; Xuyen & Trang, 2021).

The positive outcomes of this study, particularly the increased learner motivation, stem from the authentic and communicative nature of task-based activities. By engaging learners in meaningful, context-relevant communication and reducing the pressure of grammatical accuracy, TBLT fosters a supportive environment that enhances fluency and sustains motivation through greater confidence and participation. These findings corroborate the observations made by Hasibuan (2019), who identified a relationship between students' motivation and their speaking proficiency. However, they contrast with the results of Widayanti et al. (2020), which indicated that despite high speaking proficiency among students in the English Teacher Education Department, their motivation levels were low, leading to the conclusion that no positive correlation existed between speaking motivation and speaking ability. The discrepancies in findings may stem from the diverse teaching methodologies employed, the varying personality traits of the learners, and contextual factors such as the students' experiences, physical environment, and the goals behind speaking practice.

Conclusion

This study has illustrated the substantial positive effects of TBLT, particularly through form-focused tasks, on the speaking fluency and motivation of Iranian EFL students. The findings underscored that TBLT offers an engaging and effective method for enhancing speaking abilities by enabling students to practice language in authentic contexts, thereby increasing both their fluency and confidence in communication. Alongside the improvement in speaking fluency, the study also noted a significant rise in student motivation. TBLT fostered an interactive and dynamic classroom atmosphere where students could meaningfully engage with the language, moving away from conventional grammar-centric instruction. The incorporation of role-playing and communicative tasks not only advanced speaking skills but also cultivated an environment of active participation, enhancing intrinsic motivation.

Moreover, the study suggests that TBLT not only improves speaking fluency but also encourages learner autonomy. Students are motivated to take charge of their learning as they engage in tasks that necessitate real-time communication, fostering a sense of responsibility and enhancing their independent learning capabilities. This approach also promotes greater inclusivity in the classroom, allowing all students, regardless of their proficiency levels, to contribute and participate in a collaborative learning environment.

The findings of this study highlight the effectiveness of Task-Based Language Teaching (TBLT) in enhancing both speaking fluency and learner motivation among EFL students. For language educators, incorporating authentic, communicative tasks—such as role-plays and problem-solving activities—can create an engaging, interactive classroom environment that reduces speaking anxiety, fosters learner autonomy, and encourages active participation across proficiency levels. For curriculum designers, integrating meaningful, real-world tasks into course materials ensures that learners develop communicative competence alongside grammatical accuracy, while task sequencing can support gradual skill development and sustained motivation. Policymakers and institutional administrators are encouraged to support curriculum reforms that embed TBLT principles, provide professional development for teachers, and align assessment practices with communicative outcomes rather than solely grammar-focused measures.

Overall, TBLT offers a learner-centered, motivation-enhancing approach that bridges the gap between language instruction and real-life communication, suggesting that task-based activities should be a core component of effective EFL programs.

This study was limited by its short intervention, a small sample size, unexamined gender differences, and the lack of control over participants' and teachers' attitudes and types of motivation, which may have influenced the outcomes. Future research should include larger and more diverse samples, examine learners of different ages and proficiency levels, and explore the effects of TBLT on other language skills such as listening, reading, and writing. Investigating short- and long-term effects, individual differences and teachers' and students' attitudes could provide deeper insights.

Acknowledgments

This research was carried out by contribution of Omid Yeganeh, Dr. Simin Anbarshahi, Dr. Abolfazl Ghasemzadeh at Azarbaijan Shahid Madani University.

References

Afshar, H. S., & Asakereh, A. (2016). Speaking skills problems encountered by Iranian EFL freshmen and seniors from their own and their English instructors' perspectives. *Electronic journal of Foreign Language Teaching*, *13*(1), 112-130.

Ahmadian, M. J., & Long, M. H. (2020). *The Cambridge handbook of task-based language teaching.* Cambridge University Press.

Akbari, Z. (2015). Current challenges in teaching/learning English for EFL learners: The case of junior high school and high school. *Procedia - Social and Behavioral Sciences, 199*, 394–401. https://doi.org/10.1016/j.sbspro.2015.07.524

Al-Hoorie, A. H., & Dörnyei, Z. (2017). *Motivation and vision: Theoretical and empirical explorations in language learning*. Routledge. https://doi.org/10.4324/9781315200880

Al-Hosni, S. (2014). Speaking difficulties encountered by young EFL learners. *International Journal on Studies in English Language and Literature*, *2*(6), 22-30. https://www.arcjournals.org/pdfs/ijsell/v2-i6/4.pdf

Alrabai, F., & Moskovsky, C. (2016). The relationship between learners' affective variables and second language achievement. *Arab World English Journal, 7*(2), 77-103. https://doi.org/10.24093/awej/vol7no2.6

Ballesteros Soria, N., & van Compernolle (2023). Collectivizing an orientation to turn-allocation as a learnable through pre-task planning. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, *11*(2), 133-160. https://doi.org/10.22049/jalda.2024.28178.1502

Celce-Murcia, M. (2001). *Teaching English as a second or foreign language* (3rd ed.), Heinle & Heinle Publisher,

Derakhshan, A., Tahery, F., & Mirarab, N. (2015). Helping adult and young learners to communicate in speaking classes with confidence. *Mediterranean Journal of Social Sciences*, *6*(2), 520- 525. https://doi.org/10.5901/mjss.2015.v6n2p520

Detken, M. B., Ramírez-Ávila, M. R., & Chancay, C. H. (2024). Task-based activities to improve oral communication skills in adult learners. *Journal of Foreign Language Teaching and Learning, 9*(1). https://doi.org/10.18196/ftl.v9i1.19381

Dörnyei, Z. (1998). Motivation in second and foreign language learning. *Language Teaching*, *31*(3), 117-135. https://doi.org/10.1017/S026144480001315X

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Lawrence Erlbaum Associates Publishers.

Dörnyei, Z., & Otto, I., (1998), Motivation in action: A process model of L2 motivation. *Applied Linguistics, 4*, 43-69.

Dörnyei, Z., & Ushioda, E. (2011). Teaching and researching motivation (2nd ed.). Routledge. https://doi.org/10.4324/9781315833750

Dörnyei, Z., Dewaele, J.-M. (2002). *Questionnaires in second language research: Construction, administration, and processing.* Lawrence Erlbaum Associates Publishers.

Doughty, C., &Williams, J. (1998). *Focus on form in classroom second language acquisition.* Cambridge University Press.

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

Ellis, R. (2005). Principles of instructed language learning. *System*, *33*(2), 209-224. https://doi.org/10.1016/j.system.2004.12.006

Ellis, R. (2017). Position paper: Moving task-based language teaching forward. *Language Teaching, 50*(4), 507–526. https://doi.org/10.1017/S0261444817000179

Ellis, R., & Barkhuizen, G. (2005). *Analyzing learner language*. Oxford University Press.

Hashemifardnia, A., Rasooyar, H., & Sepehri, M. (2019). Enhancing Iranian EFL learners' speaking fluency through using task-based activities. *Global Journal of Foreign Language Teaching*, *9*(1), 24–32. https://doi.org/10.18844/gjflt.v9i1.3998

Hasibuan, M. S. P. (2019). *Human Resource Management*. Bandung: Citapustaka Media.

Henry, A., & Thorsen, C. (2018). Teacher–student relationships and L2 motivation. *The Modern Language Journal, 102*(1), 218–241. https://doi.org/10.1111/modl.12446

Hismanoglu, M., & Hismanoglu, S. (2011). Task-based language teaching: What every EFL teacher should do. *Procedia - Social and Behavioral Sciences 15*(3), 46-52. https://doi.org/10.1016/j.sbspro.2011.03.049

James, P. R. A. P., Yong, K. L., & Yunus, M. M. (2019). Hear me out! Digital storytelling to enhance speaking skills. *International Journal of Academic Research in Business and Social Sciences*. 9(2), 190–202. http://dx.doi.org/10.6007/IJARBSS/v9-i2/5533

Keller, J. M. (1983). Motivational design of instruction. In C. M. Reigeluth (Ed.), *Instructional-Design Theories and Models: An overview of their current status* (pp. 383–434). Lawrence Erlbaum.

Kholbutayeva, S. B. (2025). The effectiveness of task-based language teaching in enhancing speaking fluency among ESL learners. *American Journal of Social Science, 3*(3). https://doi.org/10.5281/zenodo.15062335

Kumaravadivelu, B. (2006). TESOL Methods: Changing Tracks, Challenging Trends. *TESOL Quarterly*, *40*(1), 59-81. https://doi.org/10.2307/40264511

Lam, S. F., & Law, Y. K. (2007). The roles of instructional practices and motivation in writing performance. *Journal of Experimental Education*, *75*(2), 145–164. https://doi.org/10.3200/JEXE.75.2.145-164

Lightbown, P. M. (2000). Classroom SLA research and second language teaching. *Annual Review of Applied Linguistics, 20*, 243–259. https://doi.org/10.1017/S0267190500200163

Littlewood, W. (1984). *Foreign and second language learning*. Cambridge University Press.

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (2011). 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

Mahdavirad, F. (2017). Task-Based Language Teaching in Iran: A study of EFL teachers' perspectives. *International Journal of English Language & Translation Studies. 5*(4). 14-21. https://www.eltsjournal.org/archive/value5%20issue4/3-5-4-17.pdf

Masrom, U. K., Daud, N. S. M., & Alwi, N. A. N. M. (2015). Task complexity and the complexity of written language production: A review of literature. *International Journal of Linguistics*, *7*(3), 78-93. https://doi.org/10.5296/ijl.v7i3.7553

Mystkowska-Wiertelak, A. (2011). *Review of the book* "*Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative context"*, by H. Nassaji & S. Fotos (Eds.)]. *Studies in Second Language Learning and Teaching*, *1*(3), 453–456. https://doi.org/10.14746/ssllt.2011.1.3.8

Naseri, F., Roohani, A., & Mirzaei, A. (2025). Developing L2 speaking fluency in online learning through two motivation-based instructional frameworks. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, *13*(2), 181-202. https://doi.org/10.22049/jalda.2025.30598.1801

Nget, S., Pansri, O., & Poohongthong, C. (2020). The effect of Task-Based instruction in improving English speaking skills of ninth-graders. *Language Education and Acquisition Research Network Journal*, *13*(2), 208-224.

Nita, A., Rozimela, Y., Ratmanida. (2019). The Influence of Task-Based Language Teaching on speaking skill of EFL students with intrinsic motivation. *International Journal of Literature and Arts*, *7*(6), 179-184. https://doi.org/10.11648/j.ijla.20190706.18

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language learning*, *50*(3), 417-528. https://doi.org/10.1111/0023-8333.00136

Nugrahaeni, S. M. (2022). Task-based language teaching for enhancing students' speaking ability in junior high school. *Globish: An English-Indonesian Journal for English, Education, and Culture, 11*(2), 83. https://doi.org/10.31000/globish.v11i2.6250

Nunan, D. (1999). *Second language teaching and learning*. Cambridge University Press.

Nunan, D. (2004). *Task-Based Language Teaching*. University Press.

Ochoa-Cueva, C., Cabrera, P., Quiñónez Beltrán, A. L., Castillo, L., & Gonzalez-Torres, P. (2016). The effect of communicative activities on EFL learners' motivation: A case of students in the Amazon region of Ecuador. *Colombian Applied Linguistics Journal, 18*(2), 39. https://doi.org/10.14483/calj.v18n2.10018

Salimi, A. Khayef, E., & Amini, D. (2015). The effect of strategic and online planning on accuracy of L1 learners' written production. *International Journal of Applied Linguistics and English Literature, 4(6),* 148-153**.** http://dx.doi.org/10.7575/aiac.ijalel.v.4n.6p.148

Richards, J., & Rodgers, T. (2001). *Approaches and methods in language Teaching*. Cambridge University Press.

Shirazifard, H., Abbasian, G., & Mohseni, A. (2023). EFL writing skill instruction in the light of TBLT-synthesized collaborative dialogue: Spotting learners' achievements and teachers' perceptions. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, *11*(1), 141-160. https://doi.org/10.22049/jalda.2023.28136.1497

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics*, *17*(1), 38–62. https://doi.org/10.1093/applin/17.1.38

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

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*(4), 510-532. https://doi.org/10.1093/applin/amp047

Staab, C. (1992). *Oral language for today's classroom*. Pippin Publishing.

Tale, S. M., & Goodarzi, A. (2015). The impacts of Task-Based teaching on grammar learning by Iranian first grade high school students. *International Journal of Applied Linguistics and English Literature*, *4*(4), 144-153. https://doi.org/10.7575/aiac.ijalel.v.4n.4p.144

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239–277). John Benjamins. https://doi.org/10.1075/lllt.11.15tav

Ushioda, E. (2011). Language learning motivation, self and identity: Current theoretical perspectives. *Computer Assisted Language Learning, 24*(3), 199–210. https://doi.org/10.1080/09588221.2010.538701

Weiss, D. J. (2011). Better data from better measurements using computerized adaptive testing. *Journal of Methods and Measurement in the Social Sciences 2*(1), 1-27.

Widayanti, A., Nafisah, D. A., Munawwaroh, A., Fadilah, N., & Islamīyah, M. (2020). The correlation between students' motivation and their speaking ability. In *Proceedings of the International Conference on English Language Teaching (ICONELT 2019)* (Vol. 434, pp. 153–160). Atlantis Press. https://doi.org/10.2991/assehr.k.200427.031

Willis, D. & Willis, J. (2007). *Doing Task-Based Teaching*. Oxford: Oxford University Press.

Xuyen, V. T. M., & Trang, N. H. (2021). Effects Of Task-Based speaking activities on EFL learners' oral performance. *European Journal of English Language Teaching*, *6*(6), 17-36. https://doi.org/10.46827/ejel.v6i6.3906

Zhang, Z. (2025). A review of the impact of task-based language teaching (TBLT) on oral proficiency. *Lecture Notes in Education Psychology and Public Media, 87*(1), 83-90. https://doi.org/10.54254/2753-7048/2025.BO23942

Zúñiga, E. C., Mayorga, E., & Ruiz, N. (2023). Potential implications of task-based language teaching on developing EFL learners' oral fluency. *International Journal of Learning, Teaching and Educational Research, 22*(11), 130–149. https://doi.org/10.26803/ijlter.22.11.8

Authors' Biographies

Simin Anbarshai is a full-time faculty member in the Department of English at Azarbaijan Shahid Madadni University in Iran. She graduated in English Language teaching from Panjab University of India in 2010. She joined the Department of English, Azarbaijan Shahid Madani University, as an Assistant Professor in 2012. Since then, she has constantly been involved in teaching Discourse Analysis, English for Specific Purposes, ELL, and ELT. She has presented 25 papers in different international and national congresses and also published 5 papers in domestics and international journals in different research area of ESP and ELT. She has supervised some fifteen MA students.

Omid Yeganeh received his B.A. & M.A. degrees in English Language and Literature and English Language teaching from Azarbaijan Shahid Madani University in 2023. He was constantly trying to be involved in researches and getting more knowledge in the field of English Language Teaching during his studies at university.

Volume 14, Issue 1
Winter and Spring, 2026
pp. 53-76

Journal of Applied Linguistics and Applied Literature: Dynamics and Advances

The Interplay of Emotional Intelligence and Self-Efficacy in Predicting Emotion Regulation: Insights from EFL Learners

Ali Arabmofrad[1, *], Yeganeh Kazemi[2], Milad Alizadeh[3], and Eesa Niazi[4]

[1] *Assistant Professor of Applied Linguistics, Department of English Language and Literature, Faculty of Humanities and Social Sciences, Golestan University, Gorgan, Iran, ORCID: 0000-0003-0191-2717*
Email: a.arabmofrad@gu.ac.ir

[2]*MA in TEFL, Golestan Science and Research Branch, Islamic Azad University, Gorgan, Iran Email: teflkzy@gmail.com*

[3]*PhD Candidate in ELT, Urmia University, Urmia, Iran*
Email:milad.alizadeh.d@gmail.com

[4]*Assistant Professor of Management, Department of Management, Faculty of Humanities and Social Sciences, Golestan University, Gorgan, Iran*
ORCID: 0000-0002-8606-953X
Email: e.niazi@gu.ac.ir

Abstract

Addressing an underexplored dimension of the socio-emotional aspects of second language acquisition, this study examines how emotional intelligence (EI) and self-efficacy (SE) differentially predict emotion regulation (ER) among 256 Iranian EFL learners aged 10 to 15. Employing Structural Equation Modeling (SEM), the research investigates the interrelationships among these psychological constructs within the context of English language learning. The findings indicate a moderate yet statistically significant positive association between SE and ER, and a stronger positive relationship between EI and ER, while no significant correlation emerges between SE and EI. The proposed SEM model exhibits an excellent fit, confirming that EI serves as a more powerful predictor of ER than SE. These resultsunderscore the crucial role of emotional intelligence in facilitating learners' management of affective experiences during language learning. Consequently, the study recommends incorporating EI-oriented training within EFL instructional programs to foster learners' emotional regulation, resilience, and academic success, thereby reinforcing the complementary relationship between emotional competence and linguistic-cognitive development.

Keywords: academic performance, emotional competence, psychological factors, SEM

ARTICLE INFO

Research Article
Received: Monday, June 16, 2025
Accepted: Sunday, February 15, 2026
Published: Wesensday, April 1, 2026
Available Online: Sunday, February 15, 2026
DOI: https://doi.org/ 10.22049/jalda.2026.30570.1795

Online ISSN: 2821-0204; Print ISSN: 28208986

Introduction

Salovey and Mayer's (1990) seminal model conceptualizes Emotional Intelligence as a multifaceted skill involving the discernment of one's own and others' feelings, the management of these emotions, and their application to enhance cognitive processes. Since the late 1970s, several studies have revealed that the current educational system concentrates on rational and cognitive aspects including problem-solving and memory; therefore, scant attention is paid to emotional aspect, and its crucial role is ignored. (Nelson & Low,1999). Goleman (1995) postulated that EIis the ability to identify and regulate individuals' emotions. EI can be regarded as an influential factor for predicting success in our lives. According to Goleman (1995), EI is a more reliable indicator than intelligence quotient (IQ) in determining who is likely to exceed in a given profession or role, such as becoming a top-performing salesperson, team leader, or high-ranking executive. EI functions as social intelligence and a predictor of overall performance in certain contexts, such as job performance (Salovey and Mayer, 1990)

In parallel, the concept of self-efficacy (SE), first proposed by Bandura (1977), has become a fundamental construct in the study of human behavior. Bandura (1977) defined SE as an individuals' confidence in their ability to plan and perform the actions necessary to achieve specific goals. SE affects the ways individuals tackle difficulties, establish objectives, and maintain effort when confronted with obstacles (Bandura et al., 1999). Bandura (1977) highlighted SE as a core component of human agency, asserting that individuals need to believe in their capacity to bring about change within their chosen domains in order to succeed. SE beliefs are particularly influential in shaping motivation, confidence, and stress management, especially in challenging or high-pressure situations (Bong & Clark, 1999; Bandura, 2012).

Emotions are crucial for language learning process, significantly influencing both learners' experiences and their overall achievement. According to Pekrun et al. (2022), the learning process is linked closely with a variety of emotions both positive and negative that can profoundly influence academic performance. For instance, Derakhshan (2022) emphasized that language learners often face emotional difficulties, including feelings of anxiety and frustration, alongside positive emotions like curiosity and enjoyment. Effectively managing these emotions—referred to as Emotion Regulation (ER)—is essential for successful language acquisition (Derakhshan et al., 2022; Teng & Zhang, 2021). Thompson (1994) defines ER as the extrinsic and intrinsic processes by which individuals monitor, assess, and adjust their emotional responses in order to attain particular goals.

According to Mayer and Salovey (1997), people who are skilled in emotional regulation can more effectively use strategies that diminish their negative feelings and enhance their positive ones. Moreover, ER is recognized as a key regulatory mechanism that buffers against the adverse effects of stress and promotes psychological resilience (Myruski et al., 2018).

Although numerous studies have explored the associations among emotional intelligence (EI), self-efficacy (SE), and emotion regulation (ER), there remains a significant gap concerning how these variables interact among adolescent EFL learners with lower-intermediate proficiency. To address this gap, the present investigation examines the interrelations among these three constructs and explores the extent to which SE and EI can predict learners' ability to regulate their emotions. Accordingly, the study seeks to answer the following research questions:

1. Are there any significant relationships among Iranian EFL students' ER and EI, SE?
2. Is EFL students' SE a significant predictor of students' ER?
3. Is EFL students' EI a significant predictor of students' ER?

Review of Literature

Theoretical Background

The present research is anchored in established theoretical perspectives on emotional intelligence (EI), self-efficacy (SE), and emotion regulation (ER), all of which offer an integrative foundation for examining the psychological mechanisms underlying language learning. Emotional intelligence, as defined by Salovey and Mayer (1990), refers to the capacity to recognize, comprehend, and manage one's own and others' emotions. Goleman (1995) later expanded this notion, emphasizing EI as a vital determinant of success that extends beyond conventional measures of intelligence. Complementing this, Bandura's (1977) social cognitive theory underscores the centrality of self-efficacy—individuals' beliefs in their ability to perform tasks effectively—as a driving force for motivation and persistence in educational settings. Gross (1998) adds to this framework by conceptualizing emotion regulation as the processes through which individuals modulate emotional responses to meet personal or situational goals. Integrating these perspectives allows this study to capture the interdependent nature of EI, SE, and ER, thereby illuminating their combined influence on EFL learners' affective and academic functioning.

Emotional Intelligence

Emotional Intelligence (EI), a construct rooted in Salovey and Mayer's (1990) definition concerning the recognition and management of emotions in self and others, now occupies a central position in contemporary educational inquiry. Numerous studies have established its significant impact on both academic outcomes and social functioning (Brackett & Salovey, 2006). Goleman's (1995) influential framework expanded this conceptualization by delineating five core dimensions: emotional self-awareness, impulse control, intrinsic motivation, empathetic understanding, and interpersonal effectiveness. His theoretical proposition that EI might surpass conventional cognitive measures (e.g., IQ) in predicting success across various professional domains - including leadership effectiveness, collaborative performance, and career advancement - has stimulated considerable scholarly discussion. However, subsequent empirical investigations, such as those by Van Rooy and Viswesvaran (2004), have revealed more nuanced findings, indicating that EI's predictive validity may be context-dependent and vary across different demographic groups (e.g., Alavi et al., 2019).

Within the domain of second language acquisition, EI has been demonstrated to play a pivotal role in navigating the emotional challenges associated with learning a new language. For example, Pishghadam (2009) found that language learners possessing higher EI demonstrated greater capacity in managing language anxiety and frustration, leading to improved academic performance. Similarly, Abdolrezapour and Tavakoli (2012) demonstrated that EI positively correlates with advanced cognitive processesin EFL learners. These findings indicate that EI contributes to improved emotional well-being and promotes cognitive development and academic achievement within language learning contexts.

Self-Efficacy

The emotional response patterns of individuals are influenced by SE and emotional response patterns represents individuals' EI; therefore, SE is closely linked to EI (Mills, 2014). He defined self-efficacy as an individual's belief in their capacity to execute specific tasks successfully, proposing it as a reliable predictor of subsequent performance. Bandura (1993) pointed out that a strong sense of SE accelerates an individual's success in various ways. He noted thatself-efficacious individuals considered hard tasks as challenges that must be tackled. Based on this efficient method, individuals are interested in activities, create challenging aims, and succeed. Studies indicated that learners’ SE is a dominant predictor in their educational performance (Pajares, 1996; Schunk, 1995). Based on Bandura and Schunk (1981), learners with high academic SE more probably to take on perplexing duties. On the other hand, the learners with low academic SE are less likely to deal with challenges. (Schunk, 1990).

Within the context of English language acquisition, Wang et al. (2014) conceptualized self-efficacy (SE) as learners' confidence in their capacity to successfully complete language-related tasks based on accumulated learning experiences. Extensive research has established the crucial relationship between SE and academic performance (Rafiola et al., 2020; Sirakaya et al., 2018), with multiple studies demonstrating that students with stronger self-belief tend to achieve better learning outcomes (Montgomery & Mirenda, 2014; Versland& Erickson, 2017).Although the link between self-efficacy and success is well-established, scholarly inquiry has afforded scant attention to the interaction of SE with affective elements, notably strategies for regulating emotions, within the realm of English as a Foreign Language learning.

Emotion Regulation

Numerous longitudinal investigations have been conducted on definitions of ER. ER comprises a varied range of behavioral, physiological, and cognitive processes through which individuals manage their emotional responses (Gross & John, 2003). Cole et al. (1994) hypothesized ER as the ability to navigate ongoing experiences with an appropriate range of emotional responses which are socially acceptable and adaptable—allowing for both spontaneous emotional expression and the capacity to inhibit such responses when necessary. Teng and Zhang (2017) postulated that a wide range of strategies are applied by learners to adjust their pleasant feelings, which can lead to improved academic performance. Learners apply several techniques in order to reduce their unpleasant emotions. (Fan & Wang, 2022; Greenier et al., 2021). Based on Fathi et al. (2021), learners experience adverse effects of negative feelings on their academic achievement. Therefore, learners can benefit from ER strategies which can enhance their effectiveness in the learning environment (Wang et al., 2021).ER involves various strategies employed to initiate, inhibit, or modify an individual's emotional state or behavior in specific situations (Gross, 1998). Wang and Saudino (2011) stated that ER primarily involves adjusting internal emotional states in response to external demands.

Empirical Studies

Emotional Intelligence and Self-efficacy

A robust body of research within the English as a Foreign Language (EFL) context has established a significant correlation between learners' emotional intelligence (EI) and their self-efficacy (SE). Adeyemo's (2007) study established that EFL learners with higher emotional intelligence tend to exhibit greater self-assurance, personal adequacy, and stronger beliefs in their language learning capabilities. These findings were corroborated by Hashemi and Ghanizadeh (2011), who found statistically significant relationships between EI components and EFL learners' SE beliefs. Their research particularly highlighted how emotional intelligence factors like stress tolerance and self-actualization capacity significantly correlate with learners' confidence in their language acquisition abilities. More recently, Rashid et al. (2021) further confirmed these relationships, demonstrating that various dimensions of emotional intelligence substantially contribute to the development of academic self-efficacy among language learners. Collectively, these studies suggest that emotionally intelligent learners tend to develop stronger confidence in their capacity to master a new language.

The accumulated research evidence suggests that self-efficacy (SE) influences emotional intelligence (EI) across both individual and social dimensions, while also serving as an accelerator for learning processes. Empirical investigations by Chan (2008) and Mikolajczak et al. (2007) have systematically shown that learners possessing elevated emotional intelligence typically demonstrate stronger self-efficacy beliefs and greater capacity to navigate academic challenges. The work of Rastegar and Memarpour (2009) in an Iranian EFL context provided empirical validation for this relationship, reporting a significant positive correlation between teachers' emotional intelligence and self-efficacy as measured by standardized scales (Baleghizadeh & Jula, 2024; Namaziandost et al., 2024), thereby underscoring the dynamic interplay between affective and self-belief systems in education.

Sun and Lyu (2022) investigated the relationship between university students' emotional intelligence (EI) and self-efficacy (SE), exploring the mediating role of coping styles. The participant included 800 university students in China and they were asked to complete valid questionnaire and the result analyzed by SPSS 20.0. The data revealed that SE is influenced by EI through coping styles not only directly but also indirectly. Furthermore, the results indicated that students understand more positive feelings when they measure their SE based on their emotional and mental state; therefore, they can gain higher SE in activities like occupation, education, and apprenticeships. Webb-Williams (2017) investigated science SE within primary classroom settings to identify its underlying sources. The participants of the present study were 182 children, 10 to 12 years old, and mixed-method research was conducted. The results revealed that although boys and girls achieved comparable academic performance in science, girls consistently undervalued their own capabilities. The results illustrated that gender have same academic level in science, but girls underestimate their own ability.

Emotional Intelligence and Emotion Regulation

While Rashid et al. (2021) established significant interrelationships among emotional intelligence, regulation, and expressivity in organizational contexts, a subsequent educational study by Gao and Yang (2023) delved into the predictive nature of these constructs. Surveying 391 EFL learners, they demonstrated that trait EI significantly predicts emotion regulation strategy use, with the intensity of negative emotions mediating the link to cognitive reappraisal but not to expressive suppression.

The existing literature consistently supports a robust connection between EI and ER strategies. Quintana-Orts et al. (2020) established positive correlations between these variables, while Śmieja-Nęcka and Kobylińska (2011) found that individuals with higher EI tend to employ more adaptive ER strategies. Peña-Sarrionandia et al. (2015) further elucidated this relationship through their examination of emotional processes, demonstrating that heightened EI not only influences emotional experience but also enhances regulatory capacity.

In language learning contexts, empirical evidence suggests EI's broader cognitive benefits. Abdolrezapour and Tavakoli (2012) revealed that learners with superior EI demonstrate enhanced cognitive functioning and reading abilities. Complementing these findings, Aki's (2006) research indicated that EI development can positively impact overall language acquisition outcomes.

Emotion Regulation and Self-Efficacy

The link between emotion regulation (ER) and self-efficacy (SE) is being increasingly substantiated across diverse educational populations. Initial evidence from Bazadough and Abu Sulaiman (2023) confirmed a significant correlation among university academic leaders. Building on this foundation, Doménech et al. (2024) advanced the inquiry by scrutinizing this interaction in a large adolescent cohort (N=703), introducing the moderating influence of personality traits, namely emotional stability and extraversion. Their structural equation modeling results revealed direct associations between ER and emotional SE, particularly in managing both positive and negative affective states.

In language education research, Fathi and Derakhshan (2019) focused on Iranian English instructors, utilizing standardized measures to assess how teacher SE and ER influence occupational stress. Their findings positioned SE as a stronger predictor of teaching stress compared to ER. Similarly, Chen and Lin (2009) demonstrated the predictive value of SE in academic contexts, showing its significant correlation with writing performance among Taiwanese university students completing English proficiency assessments.

Method

Participants

A purposive sampling approach was employed to select participants who met predefined inclusion criteria aligned with the study's objectives. The final sample consisted of 256 Iranian EFL learners (136 females and 120 males) aged between 10 and 15 years, all of whom were native speakers of Persian. These participants were drawn from an initial pool of 290 students who completed the Oxford Quick Placement Test (OQPT) at Milad Language Institute in Maragheh, East Azerbaijan Province, Iran. Only learners whose OQPT scores ranged between 24 and 30—indicating lower-intermediate proficiency—were included to ensure linguistic homogeneity, an essential condition for investigating the relationships among EI, SE, and ER. Informed consent was obtained from all participants or their legal guardians prior to the study, and participation was voluntary.

Instruments

Given the fact that this study was a correlational research study in essence among young intermediate learners, one homogeneity test and three types of questionnaires were administered to collect data.

Oxford Quick Placement Test

The Oxford Quick Placement Test (OQPT) is a widely recognized standardized assessment aimed at efficiently and accurately measuring the English proficiency of non-native speakers. Developed by Oxford University Press, the test assesses reading, vocabulary, grammar, and listening abilities, providing a reliable measure of a learner's CEFR (Common European Framework of Reference) level, ranging from A1 (beginner) to C2 (proficient). The first part of the test (40 questions) was administered to the students and those whose test score were between 24-30 (lower-intermediate) were selected.

Self-Efficacy Scale

This study utilized Sherer et al.'s (1982) Self-Efficacy Scale, a validated instrument measuring individuals' perceived self-efficacy across different domains. The 23-item scale features two distinct subscales: a 17-item General Self-Efficacy subscale evaluating overall confidence in handling diverse challenges, and a 6-item Social Self-Efficacy subscale assessing interpersonal competence. Participants responded using a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

Emotional Quotient Scale

Emotional intelligence was evaluated using Wong and Law's (2002) Emotional Quotient Scale, a 24-item instrument based on the four-branch EI model. The questionnaire measures four key dimensions: Self-Emotion Appraisal (recognizing one's emotions), Regulation of Emotions (managing emotional responses), Use of Emotion (harnessing emotions productively), and Others-Emotion Appraisal (understanding others' emotions). Responses were collected on a five-point Likert scale identical to the self-efficacy measure.

Emotional Regulation Questionnaire

Participants' emotion regulation strategies were assessed through Gross and John's (2003) 10-item questionnaire. This instrument distinguishes between two fundamental regulation approaches: cognitive reappraisal (adaptively reframing emotional experiences) and expressive suppression (inhibiting emotional expression). While reappraisal is associated with positive psychological outcomes, suppression often correlates with adverse effects. The instrument employs a seven-point Likert response format ranging from 1 (strongly disagree) to 7 (strongly agree).

Procedure

The study was conducted in Winter 2024. To achieve the objectives of the study participants were required to complete Oxford English proficiency test, to ensure homogeneity in language proficiency levels across the sample. They also requested to completeSEscale questionnaire, emotional quotient scale questionnaire, along with ER questionnaire. Participants were informed of the study's objectives and were encouraged to respond to the questionnaires truthfully. They were informed how to fill out the questionnaires and explained that their identities were guaranteed anonymity, after that the consent form was delivered to them. The questionnaires were distributed only to those who were willing to take part in this study after that they signed the consent form. Participants were also asked to fill out the demographic part of the questionnaire, providing information regarding their age, gender, and other details.

Data Analysis Procedure

Analysis of the collected questionnaire data was conducted using the SmartPLS software package, applying the Partial Least Squares Structural Equation Modeling (PLS-SEM) technique. The analytical procedure commenced with generating descriptive statistics to characterize the data distribution and central tendencies of the primary variables. The analytical procedure followed a two-stage approach. First, the measurement model's psychometric properties were assessed using composite reliability (CR) for internal consistency, average variance extracted (AVE) for convergent validity, and standardized factor loadings. Following the establishment of a robust measurement model, the structural model was evaluated by analyzing the significance of path coefficients, the model's explanatory power (R^2), and its predictive relevance (Q^2) to test the hypothesized relationships. The choice of PLS-SEM was appropriate due to its effectiveness in modeling complex relationships among latent constructs and its robustness with relatively small or moderate sample sizes. Furthermore, a PLS-Predict procedure was employed to assess the model’s predictive performance in out-of-sample contexts, thereby confirming the generalizability and practical implications of the findings.

Results

Descriptive Statistics

This section presents the mean, standard deviation, variance, minimum, and maximum values for each research variable, corresponding to each research question. These results are summarized in the Table 1.

Table 1

Variable Coding and Descriptive Statistics of Research Variables

	E	SE	SE	I	EA	OE	OE	EA	R
N	49	49	49	51	51	51	51	50	51
Missing	2 2 2 2 2 2 2 2 2 02 2 2 0 20 2 0 2 0 2 20 0 202 02 00 02 20 0 0 10 20 0 0 0								
Mean	.96	.93	.98	.54	.69	.36	.89	.21	.68
SD	539	533	739	91	.22	.04	.29	.24	.06
Variance	291	284	546	825	.497	.084	.655	.542	.114
Skewness	162	379	000	.675	.740	.348	1.04	.235	.300
Std. Error Skewness	154	154	154	154	154	154	154	154	154
Kurtosis	144	229	085	.320	.535	.639	.127	.972	.038
Std. Error Kurtosis	307	307	307	306	306	306	306	307	306
Minimum	.24	.47	.00	.19	.00	.00	.00	.00	.00
Maximum	.44	.71	.00	.00	.00	.00	.00	.00	.00

Note: SE = Self Efficacy; GSE = General Self-efficacy; SSE = Social Self-Efficacy; EI = Emotional Intelligence; SEA = Self-emotions appraisal; ROE = Regulation of Emotions; UOE = Use of Emotion; OEA = Others-Emotion Appraisal; ER = Emotion Regulation

The First Research Question

The first research question asked if there are any significant relationships among Iranian EFL students' emotional intelligence, self-efficacy and emotion regulation. The results of the direct relationships among the variables are presented in Table 2. The analysis, as shown in Table 2, revealed The Spearman's rho correlation analysis reveals a weak but statistically significant positive correlation ER and SE ($r = .178$, $p = .007$) and a slightly stronger significant correlation between ER and EI ($r = .249$, $p < .001$). However, no significant relationship is found between SE and ER ($r = .052$, $p = .434$).

Table 2

Results of Direct Relationships of Research Variables

			ER Mean	SE Mean	EI Mean
Spearman's rho	**ER Mean**	Correlation Coefficient	1.000	.178**	.249**
		Sig. (2-tailed)	.	.007	.000
		N	239	225	232
	SE Mean	Correlation Coefficient	.178**	1.000	.052
		Sig. (2-tailed)	.007	.	.434
		N	225	233	226
	EI Mean	Correlation Coefficient	.249**	.052	1.000
		Sig. (2-tailed)	.000	.434	.
		N	232	226	241

The Second and Third Research Questions

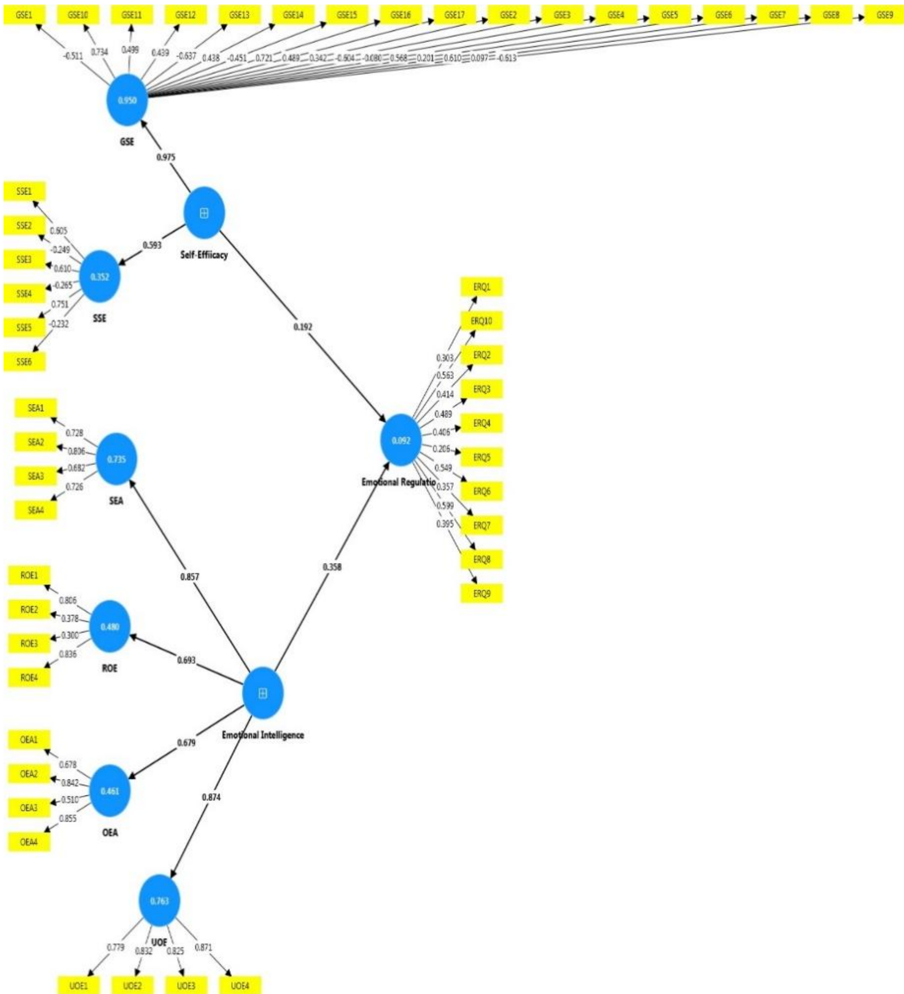
To find the predictive power of the independent variables (i.e., SE and EI) in the dependent variable (i.e., ER) the output of SmartPLS software was utilized; inferential statistics were employed to assess the questionnaire's validity and reliability, examine variable correlations, and ultimately test the research hypotheses. Model fit was assessed using outer model (measurement model) and inner model (structural model) evaluations.

Model Specification

This stage involves the formal specification of the model and is crucial in structural equation modeling (SEM). No analysis can proceed without the researcher first specifying the model, outlining the relationships between variables. This stage includes formulating an expression regarding a set of parameters. In SEM, these parameters represent the nature of the relationships between variables. SEM determines the magnitude and sign of these parameters. Figures 1 and 2 illustrate the research models, showing estimated and significant factor loadings and path coefficients. These models allow for the estimation and subsequent testing of factor loadings and path coefficients.

Figure 1

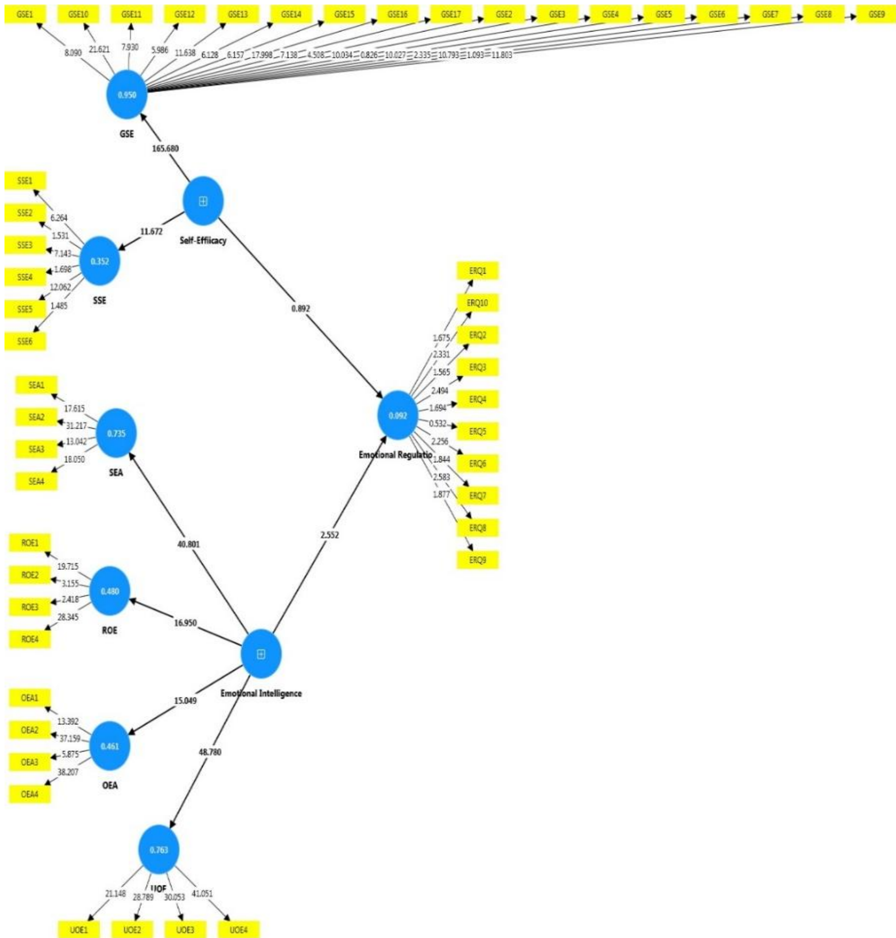
Research Model with Estimated Path Coefficients



The path diagrams feature two distinct types of coefficients. The first, termed factor loadings, represent the relationships between latent constructs and their observed indicators (rectangles). The second, path coefficients, define the structural relationships between the latent variables themselves and are used for hypothesis testing. All estimated coefficients are assessed for significance using t-statistics. Furthermore, the magnitude of a factor loading reflects an indicator's relative contribution to its underlying construct, with higher loadings denoting greater importance.

Figure 2

Research Model with Absolute Values of Significant Coefficients



Outer Model (Measurement Model) Assessment

The outer model assessment involves examining the reliability and validity of the constructs and research instruments. To assess construct reliability, three indices are used: composite reliability (CR), average variance extracted (AVE), and factor loadings (Fornell & Larcker, 1981).

Reliability Assessment

Reliability assessment in path analysis begins with the evaluation of factor loadings, which measure the strength of the relationship between latent variables and their observed indicators. A factor loading indicates how well an observed variable reflects its underlying construct. A higher loading suggests that the indicator plays a significant role in defining the construct, while a negative loading implies an inverse relationship, often due to reverse-coded items in the measurement scale. Factor loadings typically range between 0 and 1 in absolute value. According to established guidelines (Kline, 1994), loadings below 0.3 are considered weak and are usually excluded from analysis. Loadings between 0.3 and 0.6 are deemed acceptable, though they may benefit from further refinement. Loadings exceeding 0.6 are highly desirable, as they indicate a strong association between the indicator and the construct. These criteria help researchers assess whether their measurement model is reliable and whether the observed variables effectively represent the intended latent constructs.

As shown in Table 3, all factor loadings exceed 0.5, indicating that the measurement items possess adequate validity. This satisfies the first condition for establishing reliability. The third criterion for evaluating reliability involves assessing the CR of each construct. Following Fornell and Larcker's (1981) guidelines, construct reliability is determined using three key metrics: CR, AVE, and factor loadings. For a construct to be considered reliable, the CR value should exceed 0.7, while the AVE should be above 0.5. However, MacKenzie et al. (1996) suggest that an AVE of at least 0.4 may also be acceptable in certain cases.

Table 3

Factor Loadings of Indicators for Each Variable

Variable	Item	Factor Loading	Variable	Item	Factor Loading
GSE	GSE1	0.511	**SSE**	SSE1	0.605
	GSE2	0.342		SSE2	0.249
	GSE3	0.604		SSE3	0.61
	GSE4	0.08		SSE4	0.265
	GSE5	0.568		SSE5	0.751
	GSE6	0.201		SSE6	0.232
	GSE7	0.61	**SEA**	SEA1	0.728
	GSE8	0.097		SEA2	0.806
	GSE9	0.613		SEA3	0.682
	GSE10	0.734		SEA4	0.726
	GSE11	0.499	**ROE**	ROE1	0.806
	GSE12	0.439		ROE2	0.378
	GSE13	0.637		ROE3	0.3
	GSE14	0.438		ROE4	0.836
	GSE15	0.451	**UOE**	UOE1	0.779
	GSE16	0.721		UOE2	0.832
	GSE17	0.489		UOE3	0.825
ER	ERQ1	0.303		UOE4	0.871
	ERQ2	0.414	**OEA**	OEA	0.678
	ERQ3	0.489		OEA	0.842
	ERQ4	0.406		OEA	0.51
	ERQ5	0.206		OEA	0.855
	ERQ6	0.549			
	ERQ7	0.357			
	ERQ8	0.599			
	ERQ9	0.395			
	ERQ10	0.563			

For the reliability of measurement instruments, as presented in Table 4, all variables demonstrated alpha values exceeding the 0.7 threshold following validity testing, confirming satisfactory internal consistency. Additionally, the measurement model meets all established reliability criteria, with AVE values surpassing 0.5 and CR scores exceeding 0.7 for every construct. These results collectively demonstrate that the research instrument possesses strong psychometric properties and meets the required standards for reliable measurement.

Table 4
Construct Reliability Assessment

Variable	Cronbach's Alpha	CR	AVE
ER	0.851	0.829	0.738
EI	0.855	0.883	0.791
SE	0.844	0.837	0.728

Validity Assessment

Convergent Validity

Convergent validity assesses the extent to which each construct correlates with its own questions (indicators). To calculate convergent validity, AVE is used. A minimum AVE of 0.5 indicates adequate convergent validity. Table 2 shows the AVE for the constructs. Since the AVE for the model's constructs is greater than 0.5, this confirms the model's adequate convergent validity.

Discriminant Validity

The Fornell-Larcker criterion was used to assess Discriminant validity (Table 5).

	EI	**ER**	**GSE**	**OEA**	**ROE**	**SEA**	**SSE**	**SE**	**UOE**
EI	0.581								
ER	0.255	0.444							
GSE	-0.538	-0.017	0.50						
OEA	0.679	0.196	-0.32	0.73					
ROE	0.693	0.252	-0.30	0.32	0.69				
SEA	0.857	0.166	-0.41	0.47	0.50	0.93			
SSE	-0.272	0.058	0.39	-0.20	-0.16	-0.21	0.49		
SE	-0.534	0.001	0.97	-0.33	-0.29	-0.41	0.59	0.65	
UOE	0.874	0.212	-0.57	0.41	0.51	0.65	-0.2	-0.55	0.89

According to this criterion, a latent variable should have greater variance with its own observed variables than with other latent variables to demonstrate high discriminant validity. Discriminant validity of the measurement model is confirmed if the diagonal values (typically the square root of AVE) for each latent variable are greater than its correlations with other latent variables in the model.

Inner Model (Structural Model) Assessment

The assessment of structural model validity relies significantly on examining the coefficient of determination (R^2) for dependent latent variables. Chin (1998) described R^2 values of 0.19, 0.33, and 0.67 in PLS path modeling as weak, moderate, and substantial, respectively. The obtained R^2 values indicate acceptable explanatory power for the research's dependent latent variables. Unlike covariance-based approaches that use numerous model fit indices, the PLS approach lacks a "chi-square"-based fit index to examine the congruence between the theoretical model and the collected data. This is due to the predictive nature of PLS. The Q^2 statistic can be used to assess the predictive relevance of each dependent variable. A positive Q^2 value indicates adequate predictive relevance (Amani, Khezri Azar, & Mahmoudi, 2012). If the Q^2 value is positive for all dependent variables, their average can be used as an indicator of the overall quality of the structural model (Azar et al., 2012). As shown in Tables 4 and 5, the present analysis meets these criteria, with all dependent variables exhibiting positive Q^2 values.

Unlike covariance-based structural equation modeling that employs multiple fit indices, the PLS approach – being prediction-oriented – does not utilize Chi-square-based goodness-of-fit measures to evaluate theoretical model congruence with empirical data. Instead, researchers examine predictive relevance through the Q^2 statistic (Stone-Geisser criterion). As established by Amani et al. (2012), any positive Q^2 value confirms a model's predictive adequacy for its corresponding endogenous construct. Furthermore, when all dependent variables demonstrate positive Q^2 values (as shown in Table 6), their collective average serves as a comprehensive indicator of the structural model's overall predictive quality (Azar et al., 2012).

Table 6
Indices of Latent Variables

	R-square	**Q^2**
ER	0.892	0.084
GSE	0.950	0.150
OEA	0.461	0.259
ROE	0.480	0.178
SEA	0.735	0.134
SSE	0.352	0.049
UOE	0.763	0.262

For the overall model, the average communality is 0.748, and the average R^2 is 0.892. Based on the formula, the overall model fit (GoF) is 0.817, indicating a strong model fit.

Predictive accuracy was assessed using key metrics, including the coefficient of determination (R^2). The obtained R^2 value of 0.892 reveals that the model accounts for 89.2% of the variance in the outcome construct. Aligned with

Chin's (1998) benchmarks, this substantial result attests to the model's robust explanatory power and validates the postulated relationships within the theoretical framework.

Second, the model's predictive relevance was assessed through the Q^2 index, derived using the blindfolding procedure in SmartPLS software. The obtained Q^2 value of 0.284 exceeds the minimum threshold of zero required for predictive relevance (Hair et al., 2017) and surpasses the 0.25 benchmark indicating moderate predictive power. These results collectively demonstrate that the proposed model not only explains a significant portion of variance in the dependent variable but also exhibits adequate predictive capability for future observations. The combination of strong explanatory power (R^2) and satisfactory predictive relevance (Q^2) suggests that the theoretical model is both statistically robust and practically meaningful for the research context.

To strengthen the robustness of our findings, we performed PLS-Predict analysis to evaluate the model's out-of-sample predictive power. The comparative analysis of prediction errors revealed consistent superiority of our partial least squares (PLS) model over the benchmark linear model (LM). Specifically, for all measured indicators, the PLS model demonstrated lower root mean square error (RMSE) values compared to the linear model counterpart. This pattern of results was similarly observed for mean absolute error (MAE) metrics.

These findings provide compelling evidence for the model's strong predictive performance. The consistently smaller prediction errors in the PLS model suggest that our specified independent variables not only account for substantial variance in the dependent construct but also generate more accurate out-of-sample predictions than conventional linear modeling approaches. This enhanced predictive capability further validates the theoretical and practical relevance of our research model.

Discussion

The objective of this research was to delineate the relationships among EI, SE, and ER in a sample of Iranian EFL learners and to assess the predictive capacity of EI and SE on ER. The analysis revealed a significant positive relationship between EI and ER, thereby corroborating previous scholarly work (Gross, 2015; Mayer & Salovey, 1997). This result implies that heightened emotional intelligence equips learners with a greater facility for effective emotion regulation, often manifested through the use of strategies such as cognitive reappraisal (Peña-Sarrionandia et al., 2015). The strong predictive role of EI in ER underscores its importance in language learning, where emotional challenges like anxiety and frustration are common (Derakhshan, 2022).

The weak but significant SE-ER correlation ($r = .178$, $p = .007$) contradicts Bandura's (1997) assertion that self-efficacy directly enhances emotional control. This discrepancy may reflect developmental and cultural nuances. The participants (ages 10–15) likely lack the metacognitive maturity to translate SE beliefs into consistent ER strategies (Schunk, 1995). Additionally, Iran's education system, which prioritizes rote memorization over socioemotional learning (Pishghadam, 2009), may dilute SE's impact on ER. Critics might argue that the study's SE measure (general rather than language-specific efficacy) obscured domain-specific effects. Yet, even if tailored to language tasks, SE's modest role here suggests that confidence alone cannot override emotional hurdles without targeted ER instruction. This challenges the "more confidence equals better outcomes" mantra pervasive in education policy.

The non-significant SE-EI correlation defies literature positing their interdependence (Hashemi &Ghanizadeh, 2011). One explanation is compensatory overconfidence: learners with lower EI may inflate SE beliefs to cope with emotional gaps, while high-EI learners critically self-assess, lowering SE scores. Alternatively, cultural pressures to perform academically might decouple SE from emotional awareness (Wang et al., 2014). This dissonance implies that interventions boosting SE (e.g., praise-based feedback) could backfire if not paired with EI development. For instance, a learner confident in grammar drills (high SE) but unable to manage frustration (low EI) may still disengage. Thus, the study rebuts the assumption that SE and EI naturally co-evolve in educational settings.

The second research question examined whether EI could predict ER. The analysis indicated that EI significantly and positively predicted ER, supporting Mayer and Salovey's (1997) model of EI as a foundational skill for emotional management. This relationship aligns with previous research showing that individuals with higher EI are more likely to use adaptive regulation strategies, such as cognitive reappraisal (Gross, 2015). Within the EFL context, where learners oftenencounter anxiety and frustration (Derakhshan, 2022), EI components such as emotional awareness and understanding appear crucial for effective emotion modulation. However, the moderate strength of this relationship suggests EI operates alongside other factors, potentially including classroom climate or personality traits (Solhi et al., 2023). The findings strongly support incorporating EI development into language curricula, particularly in contexts like Iran where traditional education may neglect emotional skills (Pishghadam, 2009). Future interventions should focus on cultivating specific EI competencies that directly support emotion regulation in language learning situations.

Although the current findings are consistent with earlier research that underscores the influential role of emotional intelligence (EI) in enhancing emotion regulation (ER), this study contributes a distinct perspective by situating these constructs within the socio-cultural and developmental context of adolescent Iranian EFL learners. The results indicate that EI serves as a stronger predictor of ER than self-efficacy (SE), suggesting that emotional competencies play a particularly salient role in coping with the affective challenges of language learning—an aspect often underemphasized in cognitively oriented educational environments. Contrary to Bandura’s (1997) proposition that SE directly shapes emotional control, the modest relationship observed here implies that self-efficacy may exert only an indirect influence on ER. Consequently, pedagogical interventions designed to foster emotion regulation might benefit from incorporating contextually sensitive strategies that strengthen both learners' emotional and efficacy beliefs. By interpreting the results through this culturally informed lens, the present discussion extends previous findings and offers practical implications for EFL instruction and emotional skills training.

The results of this study carry important implications for language learning and teaching. First, the demonstrated significant relationship between EI and ERhighlights the necessity of incorporating emotional skills training into language curricula. By teaching learners how to recognize, understand, and regulate their emotions, educators can help them manage the emotional difficulties of language learning more effectively. This, in turn, could lead to enhanced academic achievement and greater overall satisfaction with the learning process.

Conclusion

In summary, this research contributes to the extant body of knowledge by delineating the relationships among Emotional Intelligence (EI), self-efficacy (SE), and emotion regulation (ER) within an Iranian EFL context. The findings foreground the fundamental importance of EI in underpinning adaptive ER, positing that pedagogical interventions aimed at cultivating emotional competencies could significantly improve learners' management of affective demands. Conversely, the absence of a significant SE-ER relationship challenges prevailing assumptions and highlights the imperative for a more refined, culturally-situated theoretical model. Such scholarly refinement is a prerequisite for designing empirically-grounded and pedagogically responsive learning atmospheres that concurrently facilitate emotional and educational development.

The study also calls for further research to explore the contextual and cultural factors that may influence these relationships. In conclusion, while EI emerges as a cornerstone of emotion regulation in language learning, the role of SE appears context-dependent. These findings advocate for a paradigm shift in EFL pedagogy, prioritizing emotional literacy alongside linguistic competence to cultivate resilient, adaptive learners.

Acknowledgments: We extend our sincere gratitude to the EFL learners from Milad Language Institute for their valuable time and contributions.

References

Abdolrezapour, P., & Tavakoli, M. (2012). The relationship between emotional intelligence and EFL learners' achievement in reading comprehension. *Innovation in Language Learning and Teaching, 6*(1), 1–13. https://doi.org/10.1080/17501229.2010.550686

Adeyemo, D. A. (2007). Moderating influence of emotional intelligence on the link between academic self-efficacy and achievement of university students. *Psychology and Developing Societies, 19*(2), 199–213. https://doi.org/10.1177/097133360701900204

Alavi, M., Mehrinezhad, S. A., Amini, D., Ningal, M. T., & Abdul Latif, A. (2019). Comparative study of trait emotional intelligence and executive functioning among youth. *The Journal of General Psychology, 146*(1), 50-67. https://doi.org/10.1080/00221309.2018.1535484

Baleghizadeh, S. and Jula, A. (2024). Sense of self-efficacy and emotional intelligence as predictors of job satisfaction among Iranian EFL teachers. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, *12*(2), 161-194. https://doi.org/10.22049/jalda.2024.29316.1645

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*(2), 117–148. https://doi.org/10.1207/s15326985ep2802_3

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management, 38*(1), 9–44. https://doi.org/10.1177/0149206311410606

Bandura, A., & Schunk, D.H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology, 41*(3), 586–598. http://dx.doi.org/10.1037/0022-3514.41.3.586

Bandura, A., Freeman, W.H., & Lightsey, R. (1999). *Self-efficacy: The exercise of control*. Berlin: Springer. https://doi.org/10.1891/0889-8391.13.2.158

Bazadough, D. S., & Abu Sulaiman, B. (2023). Emotional regulation and its relationship to self-efficacy among academic leaders in Amman Arab University. *Journal of Namibian Studies: History Politics Culture, 33*(1), 69-85, *Article 565*. https://doi.org/10.59670/jns.v33i.565

Bong, M., & Clark, R. E. (1999). Comparison between self-concept and self-efficacy in academic motivation research. *Educational Psychologist, 34*(3),139–153. https://doi.org/10.1207/s15326985ep3403_1

Brackett, M. A., & Salovey, P. (2006). Measuring emotional intelligence with the Mayer-Salovery-Caruso Emotional Intelligence Test (MSCEIT). *PubMed, 18(Suppl)*, 34–41.

Chan, D. W. (2008). Emotional intelligence, self-efficacy, and coping among Chinese prospective and in-service teachers in Hong Kong. *Educational Psychology, 28*(4), 397–408. https://doi.org/10.1080/01443410701668372

Chen, M. C., & Lin, H. J. (2009). Self-efficacy, foreign language anxiety as predictors of academic performance among professional program students in a general English proficiency writing test. *Perceptual & Motor Skills, 109*(2), 420–430. http://dx.doi.org/10.2466/PMS.109.2.420-430

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 295–336). Lawrence Erlbaum Associates.

Cole, P. M., Michel, M. K., & Teti, L. O. (1994). The development of emotion regulation and dysregulation: A clinical perspective. *Monographs of the Society for Research in Child Development, 59*(2/3), 73-91. https://doi.org/10.2307/1166139

Derakhshan, A. (2022). Revisiting research on positive psychology in second and foreign language education: Trends and directions. *Language Related Research, 13*(5), 1–43. https://doi.org/10.52547/LRR.13.5.2

Derakhshan, A., Fathi, J., Pawlak, M., & Kruk, M. (2022). Classroom social climate, growth language mindset, and student engagement: The mediating role of boredom in learning English as a foreign language. *Journal of Multilingual and Multicultural Development, 43,* 1–19. https://doi.org/10.1080/01434632.2022.2099407

Doménech, P., Tur-Porcar, A. M., & Mestre-Escrivá, V. (2024). Emotion regulation and self-efficacy: The mediating role of emotional stability and extraversion in adolescence. *Behavioral Sciences, 14*(3), 206-228. https://doi.org/10.3390/bs14030206

Fan, J., & Wang, Y. (2022). English as a foreign language teachers' professional success in the Chinese context: The effects of well-being and emotion regulation. *Frontiers in Psychology, 13*, 952503. https://doi.org/10.3389/fpsyg.2022.952503

Fathi, J., & Derakhshan, A. (2019). Teacher self-efficacy and emotional regulation as predictors of teaching stress: an investigation of Iranian English language teachers. *Teaching English Language*, *13*(2), 117–143.

Fathi, J., Greenier, V., & Derakhshan, A. (2021). Self-efficacy, reflection, and burnout among Iranian EFL teachers: The mediating role of emotion regulation. *Iranian Journal of Language Teaching Research, 9*(2), 13–37. https://doi.org/10.30466/ijltr.2021.121043

Gao, Z., & Yang, Y. (2023). The predictive effect of trait emotional intelligence on emotion regulation strategies: The mediating role of negative emotion intensity. *System*, *112*, *102958*. https://doi.org/10.1016/j.system.2022.102958

Goleman, D. (1995). *Emotional intelligence*. Bantam Books.

Greenier, V., Derakhshan, A., & Fathi, J. (2021). Emotion regulation and psychological well-being in teacher work engagement: A case of British and Iranian English language teachers. *System, 97*(1), 102446. https://doi.org/10.1016/j.system.2020.102446

Gross, J. J. (1998). Antecedent- and response-focused emotion regulation: divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology, 74*, 224–237. https://doi.org/10.1037//0022-3514.74.1.224

Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological Inquiry, 26*(1), 1–26. https://doi.org/10.1080/1047840X.2014.940781

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*(2), 348–362. https://doi.org/10.1037/0022-3514.85.2.348

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling* (PLS-SEM) (2nd ed.). SAGE Publications.

Hashemi, M. R., &Ghanizadeh, A. (2011). Emotional intelligence and self-efficacy: A case of Iranian EFL university students. *International Journal of Linguistics, 3*(1), 35-57. https://doi.org/10.5296/ijl.v3i1.877

Namaziandost, E., Heydarnejad, T., & Rezai, A. (2024). Teaching style preferences as a function of self-efficacy, emotion regulation, reflective teaching, and mindfulness in teaching: A voice from an EFL context. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, *12*(2), 1-29. https://doi.org/10.22049/jalda.2024.28775.1568

Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey, & D. J. Sluyter (Eds.), *Emotional development and emotional intelligence: Educational implications* (pp. 3-31). Basic Books.

Mikolajczak, M., Menil, C., & Luminet, O. (2007). Explaining the protective effect of trait emotional intelligence regarding occupational stress: Exploration of emotional labour processes. *Journal of Research in Personality, 41*(5), 1107–1117. https://doi.org/10.1016/j.jrp.2007.01.003

Mills, N. (2014). Self-efficacy in second language acquisition. In S. Mercer & M. Williams (Eds.), *Multiple perspectives on the self in SLA* (pp. 6-22). Multilingual Matters.

Montgomery, A., & Mirenda, P. (2014). Teachers' self-efficacy, sentiments, attitudes, and concerns about the inclusion of students with developmental disabilities. *Exceptionality Education International, 24*(1), 55-77. https://doi.org/10.5206/eei.v24i1.7708

Myruski, S., Denefrio, S., & Dennis-Tiwary, T. A. (2018). *Stress and emotion regulation*. Oxford Handbooks Online.

Nelson, D., & Low, G. (1999). *Exploring and developing emotional intelligence skills*. Kingsville, TX: EI Learning Systems.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543– 578. http://dx.doi.org/10.3102/00346543066004543

Pekrun, R., Marsh, H. W., Suessenbach, F., Frenzel, A. C., & Goetz, T. (2022). School grades and students' emotions: Longitudinal models of within-person reciprocal effects. *Learning and Instruction*, *83*: 101626. https://doi.org/10.1016/j.learninstruc.2022.101626

Peña-Sarrionandia, A., Mikolajczak, M., & Gross, J. J. (2015). Integrating emotion regulation and emotional intelligence traditions: A meta-analysis. *Frontiers in Psychology, 6*: 160. https://doi.org/10.3389/fpsyg.2015.00160

Pishghadam, R. (2009). A quantitative analysis of the relationship between emotional intelligence and foreign language learning. *Electronic Journal of Foreign Language Teaching. 6*(1), 31–41.

Quintana-Orts, C., Mérida-López, S., Rey, L., Neto, F., & Extremera, N. (2020). Untangling the emotional intelligence-suicidal ideation connection: The role of cognitive emotion regulation strategies in adolescents. *Journal of Clinical Medicine, 9*(10), 3116. https://doi.org/10.3390/jcm9103116

Rafiola, R. H., Setyosari, P., Radjah, C. L., & Ramli, M. (2020). The effect of learning motivation, self-efficacy, and blended learning on students' achievement in the industrial revolution 4.0. *International Journal of Emerging Technologies in Learning (IJET), 15*(8), 71-89. https://doi.org/10.3991/ijet.v15i08.12525

Rashid, A., Akhtar, M., & Riaz, M. N. (2021). Relationship between emotional intelligence, emotion regulation and emotional expressivity in employees. *International Review of Management and Business Research, 10*(1), 361–366. https://doi.org/10.30543/10-1(2021)-31

Rastegar, M., &Memarpour, S. (2009). The relationship between emotional intelligence and self-efficacy among Iranian EFL teachers. *System, 37*(4), 700–707. https://doi.org/10.1016/j.system.2009.09.013

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality, 9*(3), 185–211. https://doi.org/10.2190/DUGG-P24E-52WK-6CDG

Schunk, D. H. (1990). Goal-setting and self-efficacy during self-regulated learning. *Educational Psychologist, 25*(1), 71–86. http://dx.doi.org/10.1207/s15326985ep2501_6

Schunk, D. H. (1995). Self-efficacy and education and instruction. In J. E. Maddux (Ed.), *Self-efficacy, adaptation, and adjustment: Theory, research, and application* (pp. 281–303). Plenum Press.

Sherer, M., Maddux, J. E., Mercandante, B., Prentice-Dunn, S., Jacobs, B., & Rogers, R. W. (1982). The self-efficacy scale: Construction and validation. *Psychological Reports, 51*(2), 663–671. https://doi.org/10.2466/pr0.1982.51.2.663

Sirakaya, M., & Cakmak, E. (2018). Effects of augmented reality on student achievement and self-efficacy in vocational education and training. *International Journal for Research in Vocational Education and Training (IJRVET), 5*(1), 1–18. https://doi.org/10.13152/IJRVET.5.1.1

Śmieja-Nęcka, M., & Kobylińska, D. (2011). Emotional intelligence and emotion regulation strategies. *Studia Psychologiczne, 49*(5), 55–64. https://doi.org/10.2478/v10167-010-0040-x

Solhi, M., Derakhshan, A., & Ünsal, B. (2023). Associations between EFL students' L2 grit, boredom coping strategies, and emotion regulation strategies: a structural equation modeling approach. *Journal of Multilingual and Multicultural Development*, *51*(2), 1–20. https://doi.org/10.1080/01434632.2023.2175834

Sun, G., & Lyu, B. (2022). Relationship between emotional intelligence and self-efficacy among college students: the mediating role of coping styles. *Discover Psychology, 2*(1), 78-86. https://doi.org/10.1007/s44202-022-00055-1

Teng, L. S., & Zhang, L. J. (2017). Effects of motivational regulation strategies on writing performance: a mediation model of self-regulated learning of writing in English as a second/foreign language. *Metacognition and Learning, 13*(2), 213–240. https://doi.org/10.1007/s11409-017-9171-4

Teng, L. S., & Zhang, L. J. (2021). Can self-regulation be transferred to second/foreign language learning and teaching?: Current status, controversies, and futures Directions. *Applied Linguistics*. https://doi.org/10.1093/applin/amab032

Thompson, R. A. (1994). Emotion regulation: A theme in search of definition. *Monographs of the Society for Research in Child Development, 59*(2/3), 25–52. https://doi.org/10.2307/1166137

Van Rooy, D. L., &Viswesvaran, C. (2004). Emotional intelligence: A meta-analytic investigation of predictive validity and nomological net. *Journal of Vocational Behavior, 65*(1), 71–95. https://doi.org/10.1016/S0001-8791(03)00076-9

Versland, T. M., & Erickson, J. L. (2017). Leading by example: A case study of the influence of principal self-efficacy on collective efficacy. *Cogent Education, 4*(1), 121-125, Article 1286765. https://doi.org/10.1080/2331186x.2017.1286765

Wang, C., Kim, D.-H., Bai, R., & Hu, J. (2014). Psychometric properties of a self-efficacy scale for English language learners in China. *System, 44*(3), 24–33. http://dx.doi.org/10.1016/j.system.2014.01.015

Wang, M., & Saudino, K. J. (2011). Emotion regulation and stress. *Journal of Adult Development, 18*(2), 95–103. https://doi.org/10.1007/s10804-010-9114-7

Wang, Y., Derakhshan, A., & Zhang, L. J. (2021). Researching and practicing positive psychology in second/foreign language learning and teaching: The past, current status and future directions. *Frontiers in Psychology, 12*(12), 57-59. https://doi.org/10.3389/fpsyg.2021.731721

Webb-Williams, J. (2017). Science self-efficacy in the primary classroom: Using mixed methods to investigate sources of self-efficacy. *Research in Science Education, 48*(5), 939–961. https://doi.org/10.1007/s11165-016-9592-0

Wong, C-S., & Law, K. S. (2002). Wong and Law Emotional Intelligence Scale. *PsycTESTS Dataset.* https://doi.org/10.1037/t07398-000

Authors' Biography

Ali Arabmofrad is an Assistant Professor at Golestan University in Gorgan, Iran. He received his PhD in TEFL from University of Tehran, Iran in 2014. He has published a number of articles both in national and international journals (Reading & Writing Quarterly, English Teaching & Learning, Journal of Research in Childhood Education, Journal of Psycholinguistic Research, etc.). His areas of interest include psycholinguistics, pragmatics and first language acquisition.

Yeganeh Kazemi received her M.A. in TEFL from Golestan Science and Research Branch, Islamic Azad University, Gorgan, Iran, in 2016. She received her B.A. in English Language and Literature from Golestan University, Gorgan, Iran in 2013. She has published some articles. Her areas of interest include language learning, language learning strategies, and MALL.

Milad Alizadeh received his B.A., and M.A. degrees in English Language Teaching from Urmia University, and Maragheh University, Iran, respectively. Currently, he is a Ph.D. candidate of English Language Teaching at Urmia University, Urmia, Iran. In 2017, he founded a languages center in Urmia, Iran. Since then, he has supervised many language teachers and learners in branches of the languages Center. He has authored some articles which are under review.

Eesa Niazi is an Assistant Professor at Golestan University in Gorgan, Iran. He received his PhD in Management from University of Tehran, Iran in 2018. He has published a number of articles in national journals (Educational Leadership and Management Research, Applied Educational Leadership, Journal of Executive Management, etc.). His areas of interest include psycholinguistics, educational leadership and education policymaking.

Volume 14, Issue 1
Winter and Spring, 2026
pp. 77-98

Journal of Applied Linguistics and Applied Literature: Dynamics and Advances

A Corpus-Based Investigation of Idioms in Applied Linguistics Research Articles

Elham Farahani[1, *] and Hooshang Yazdani[2]

[1] *Corresponding Author: Assistant Professor of TEFL, Department of English Language and Literature, Faculty of Letters and Languages, Arak University, Arak, Iran; ORCID: 0000-0001-9068-755X*
Email: e-farahani@araku.ac.ir

[2] *Associate Professor of TEFL, Department of English Language and Literature, Faculty of Letters and Languages, Arak University, Arak, Iran*
ORCID: 0000-0003-3628-1046
Email: h-yazdani@araku.ac.ir

Abstract

Idioms are typically considered as elements of informal, vivid, and entertaining language, and are generally thought to be unsuitable for formal contexts. However, recent studies show that idioms are increasingly used in formal communication and academic writing. In this context, this corpus-based study aims to offer new insights into English idioms by examining them within a corpus consisting of 5,675,554 tokens from 640 randomly selected applied linguistics research articles (RAs). To achieve this, the researchers developed a working definition and framework for identifying academic idioms within written genres like RAs. Using this framework, they compiled a list of the most common academic idioms found in the genre. Additionally, the study explored whether the frequency of these idioms changed over the period from 1980 to 2020. The findings revealed that idioms are indeed present in applied linguistics RAs and merit appropriate attention. No significant correlation was found between time and idiom usage frequency, while the writing style of authors appeared to significantly influence how often idioms are used. By identifying and analyzing the most frequent academic idioms in applied linguistics RAs, this study's results may be valuable for EFL writing instructors, graduate and postgraduate students in applied linguistics, and developers of teaching materials.

Keywords: idioms, research articles, corpus, applied linguistics

ARTICLE INFO

Research Article
Received: Sunday, November 2, 2025
Accepted: Sunday, March 15, 2026
Published: Wedensday, April 1, 2026
Available Online: Sunday, March 15, 2026
DOI: https://doi.org/ 10.22049/jalda.2026.31176.1916

Online ISSN: 2821-0204; Print ISSN: 28208986

Introduction

RAs "the written form of the academy" (van Enk & Power, 2017, p.1), are the subject of numerous English for Research Publication Purposes (ERPP) and English for Academic Purposes (EAP) studies. In fact, RAs have become "the pre-eminent genre of the academy" (Hyland, 2009, p. 67). This is due to the significance of RAs in dissemination of knowledge in academic communities. In this regard, Hyland & Paltridge (2011) state: "publishing is the main means by which academics establish their claims for competence and climb the professional ladder" (p. 173). The saying: "publish or perish" accurately describes the vital role of RAs publication in the academic lives of scholars and graduate students.

English as a Foreign Language (EFL) graduate students and teachers are expected to publish RAs in English to be accepted in their academic communities. This requirement seems to be less challenging for applied linguistics graduate students since they have passed some English writing courses. However, these courses focus on general writing skills, not to RAs. Therefore, their academic writing skills usually lag behind their level of conversational skills. This is unsurprising since even native speakers of any language require formal training to write academic RAs.

As a result, EAP researchers aim to investigate the features of academic writing and the variables contributing to successful composition of RAs acceptable to members of their scientific communities. Using corpus-based methodology, this study investigates a neglected aspect: *academic idioms*.

The investigation of idioms in RAs may seem odd since they are often assumed to be more prevalent in informal communications, only oiling the wheels of friendly and conversational language (Miller, 2020; Simpson & Mendis, 2003). O'Keeffe et al. (2007, p. 90) state that an academic "context is not immediately associated with the occurrence of idiomatic expressions in most people's minds". They indicate that "idioms have functions which create and reinforce interpersonal relations, projecting informality, camaraderie and social bonding" (p. 88).

An examination of ten EAP books revealed that most of them warn learners about the use of idioms in their academic writing. For example, Sowton (2012) states "since one of the main goals of academic writing is to avoid ambiguity, idioms should be avoided." (p. 89). Likewise, Bailey (2015) in his EAP book states that "there is no correct style of academic writing, but in general it should be accurate, impersonal and objective. For example, personal pronouns like I and idioms (i.e., informal language) are used less often than in other kinds of writing" (p. 138).

What is more, in coursebooks on idioms too, academic idioms are ignored implying learners should avoid using idioms in their academic discourse. As an example, in McCarthy and O'Dell's (2010) self-study book on idioms, only one lesson is devoted to the formal idioms that can safely be used in writing.

However, it has been found that idioms which may strike an informal note in dry and academic contexts are frequent in spoken academic discourses (Miller, 2020; Simpson & Mendis, 2003). For instance, O'Keeffe et al. (2007) give examples of the use of idioms in academic situations in which people feel the necessity to make use of idioms to project a more friendly relationship to express meanings non-threateningly.

Related to this discussion is Biber et al.'s (1999) massive reference book which is a corpus-informed investigation on written and spoken English in four registers (conversation, fiction, news, and academic prose) from American and British English. Interestingly, one part of this influential book is devoted to idioms. In this part, they point out that a large number of idioms in English are verb-based (i.e., *bear in mind*, *take into account*, *stand a chance*), and that idioms are generally rare in conversation, and more common in fiction, a finding surprising those believing idioms belong to daily talk.

Biber et al. (1999) confirm that idioms are mostly colloquial. Crucially, they report that although idioms are generally colloquial and more frequent in other registers, "a few of these expressions are used 'primarily' in academic prose or other kinds of expository writing" (p.1026). They add: "although earlier studies have claimed that formulaic language is dominant in conversation, idiomatic phrases with the verbs *have*, *make*, and *take* are by far more common in the written registers. In fact, several of these phrases are notably common only in news reportage and/or academic prose." (p. 1028). Idioms like *have an/no/the effect*, *make use of*, and *take the form of* are reported to be used only in academic discourse.

Researchers (McCarthy, 1998; Moon, 1998; Simpson & Mendis, 2003) conclude that the use of idioms are not merely alternatives to their more formal and literal counterparts. In fact, idioms are found to perform discourse roles and important pragmatic functions in academic contexts.

Therefore, it stands as a reason to conclude that, although idioms in academic English lack the frequency of informal English, they have important functions in this genre. As Miller (2020) states, low frequency does not equal low importance since students may miss or misunderstand text due to unfamiliar idioms. For example, ignorance of the idiomatic meanings of *that said*, *much as,* or *nothing if not* as academic idioms can lead to total misunderstanding of the sentences. Another example is *in view of something* that literally has nothing to do with the meaning of the whole: *because of something or as a result of something*.

In this study, idioms are not restricted to the vivid and colorful examples. Many phrases qualify as idioms because their form is fixed, and/or they show distinctive grammar. For example, *for example* is regarded as an idiom in this study. Though semantically transparent, the grammar is idiomatic because it treats 'example' as if it were an uncountable noun, which in general it is not. The totally unidiomatic expression would be '*for an example*', but it is not used by English speakers. English speakers would say '*as an example*', which is quite straightforward both semantically and grammatically, and therefore it is treated not as an idiom.

In sum, it is safe to say that idiomaticity is not a binary quality – it exists on a cline. Some expressions are more idiomatic than others and deciding whether a particular expression is or is not an idiom may depend on where exactly you draw the line; this in turn may depend on what a researcher's purpose is.

Taken all these points into account, one of the primary purposes this study aims to accomplish is to challenge one of the long-held assumptions and stereotypes attested to idioms that they are not qualified enough to be used in academic writing. Moreover, attempts are made to present an operational definition and a framework for the detection of academic idioms in this corpus-informed study to develop a list of genre and discipline-specific academic idioms which are compiled from applied linguistics RAs published in eight professional journals in the field from 1980 to 2020. This list is the first attempt in the literature and is hoped to be insightful for EAP and EFL writing courses. Finally, any possible changes in idiom frequency used in applied linguistics RAs over the time span of 1980-2020 are investigated. This study aims to answer the following research questions:

1. How can academic idioms be defined to distinguish them from other multiword expressions?
2. What is the frequency of idioms in applied linguistics RAs?
3. Are there differences in idiom frequency in applied linguistics RAs from 1980 to 2020?

In order to address the third research question, the following null hypothesis is constructed to be tested out:

H01: There are no differences in idiom use in applied linguistics RAs from 1980 to 2020.

Literature Review

Definitions of Idioms

The term "idiom" derives from Greek, meaning "own, private, and peculiar" (Simpson & Weiner, 1989, p. 624), referring to phrases with unique meanings or structures (Panou, 2014). Idiomaticity describes expressions that are grammatical and native-like or non-compositional, such as *a dog's breakfast* (meaning a mess) (Richards & Schmitt, 2010). Sinclair's (1991) idiom principle posits that language users often select pre-fabricated phrases for efficiency.

Many researchers have long tried to provide a good definition for idioms. At one extreme, Hocket (1958) includes single morphemes to entire texts, while Grant and Bauer (2004) limit idioms to 140 "pure" cases. Weinreich (1969) and Fraser (1970) emphasize non-compositional meanings, and Makkai (1972) distinguishes transparent (encoding) and opaque (decoding) idioms. Nunberg (1978) categorizes idioms as decomposable (e.g., *pop the question*) or non-decomposable (e.g., *by and large*). Grant and Bauer (2004) classify idioms into core (non-compositional, non-figurative), figurative, and ONCE idioms. Gramley and Patzold (2004) highlight semantic contrast, dual meanings, and syntactic irregularities as defining traits.

Significance of Idioms

Idioms are integral to natural language, adding color and variety (Liontas, 2017; Cooper, 1998). With over 10,000 idioms in American English (Brenner, 2011), they enhance communication efficiency by replacing lengthy explanations (e.g., *hot potato*) (Khonbi & Sadeghi, 2017). Idioms contribute to fluency by reducing cognitive load through fixed phrases (Pawley & Syder, 1983, as cited in Hinkel, 2017). Idiomatic competence, developed through exposure, marks fluency and cultural integration (Liontas, 2017). However, misinterpreting idioms can lead to comprehension issues (Boers et al., 2007).

Properties of Idioms

Idioms are characterized by multi-wordiness, non-compositionality, fixedness, and institutionalization. They typically involve multiple words (Dobrovol'skij & Piirainen, 2005). Non-compositionality, where meaning isn't derived from individual words, is debated, as many idioms have metaphorical or historical origins (Boers et al., 2007). Fixedness is also contested, with corpus studies showing syntactic flexibility (Moon, 1998; Glucksberg et al., 2001). Institutionalization reflects idioms' acceptance as conventional expressions (Fernando, 1996).

Idioms vs. Other Multi-Word Expressions

Idioms differ from other multi-word expressions in their fixed, non-compositional nature, which makes their meanings less predictable. For instance, collocations, such as *strong tea*, are transparent and flexible in meaning, allowing for easier interpretation compared to idioms like *call into question*, which rely on fixed phrasing (Baker, 2018; Gramley & Patzold, 2004). Metaphors, like *guiding light*, implicitly compare distinct concepts, while similes use explicit markers such as "like" or "as" for comparison, yet both contrast with idioms' non-literal rigidity (Leah, 2014; McCarthy & O'Dell, 2008). Proverbs, such as *the nail that sticks up*, convey cultural wisdom with relatively transparent meanings, unlike the non-narrative nature of idioms (Dobrovol'skij & Piirainen, 2005). Clichés, like *have a good day*, are fixed yet transparent, setting them apart from the opacity of idioms (Gramley & Patzold, 2004).

Corpus-Based Studies on Idioms

Corpus studies illuminate idiom usage. Moon (1998) found 40% of idioms in an 18-million-word corpus showed variation, though her corpus was journalism-heavy. Simpson and Mendis (2003) identified idioms in academic speech, noting functions like evaluation and emphasis. Liu (2003) reported rare pure idioms in spoken English. Grant and Nation (2006) listed 100 core idioms, emphasizing their scarcity. Miller (2020) confirmed idioms' presence in academic writing, identifying 43 frequent idioms in the Oxford Corpus of Academic English, suggesting their relevance for RAs. Alizadeh et al. (2024) compared the distinctive use of lexical bundles in soft and hard science publications. However, no study specifically addresses idioms in applied linguistics RAs, a gap this research aims to fill.

The literature reveals diverse idiom definitions, their significance in fluency and cultural integration, and their debated properties. Corpus studies highlight idioms' variability and presence in academic discourse, but their role in applied linguistics RAs remains underexplored. This study seeks to address this gap, providing insights into idioms' academic utility.

Method

Corpus Design

To investigate idiomatic expressions in RAs within applied linguistics, a specialized corpus was developed due to the absence of an existing online corpus tailored for this purpose. The corpus comprised 640 RAs randomly selected from eight high-ranking applied linguistics journals, identified using Scimagojr.com based on impact factor. Journals were included if published from 1980 to 2020, ensuring a 40-year span. The selected journals were *Applied Linguistics*, *ELT Journal*, *Journal of Communication*, *Journal of Research Literacy*, *Language Learning*, *TESOL Quarterly*, *The Modern Language Journal*, and *The Journal of Child Language*. For each journal, two RAs per year were randomly downloaded, yielding 80 articles per journal and a total of 5,675,554 tokens.

Data Analysis Tool

The corpus was analyzed using #LancsBox (Brezina et al., 2018, 2020), a sophisticated corpus analysis tool developed by Lancaster University. #LancsBox was chosen for its advanced filtering capabilities, ability to process various file formats (e.g., .txt, .pdf, .docx), and efficient statistical analysis of large datasets. The KWIC and Whelk tools provided concordance lines, frequency, relative frequency, and distribution data across the corpus, facilitating idiom identification.

Data Analysis Procedure

Given the complexity of idioms, automatic corpus searches were impractical. Instead, a two-stage approach was adopted, inspired by Miller (2020) and Simpson and Mendis (2003).

Stage One: Manual Analysis

In the first stage, 240 RAs (30 per journal) were manually analyzed to develop an operational definition and framework for identifying academic idioms. These articles, covering 1980–2020, were read to extract multi-word expressions, which were then evaluated against idiom dictionaries and criteria from the literature. Three phraseology experts provided feedback to refine the framework. The extracted idioms (6,834) were categorized by decade (1980–1989, 1990–1999, 2000–2009, 2010–2020) and sorted by frequency using a coding system, resulting in a list of 567 idioms.

Stage Two: Corpus-Wide Search

In the second stage, the 567 idioms were searched across the full 640-article corpus using #LancsBox. Advanced filtering accounted for idiomatic variations (e.g., *along/on (the) lines*), ensuring all forms were captured. For instance, filtering for *to a/an extent* reduced 1,049 concordance lines to 101. Idioms with dual meanings (e.g., *on the ground*) were manually checked to distinguish idiomatic from literal uses. Frequency, relative frequency, and range data were recorded for each idiom.

Results

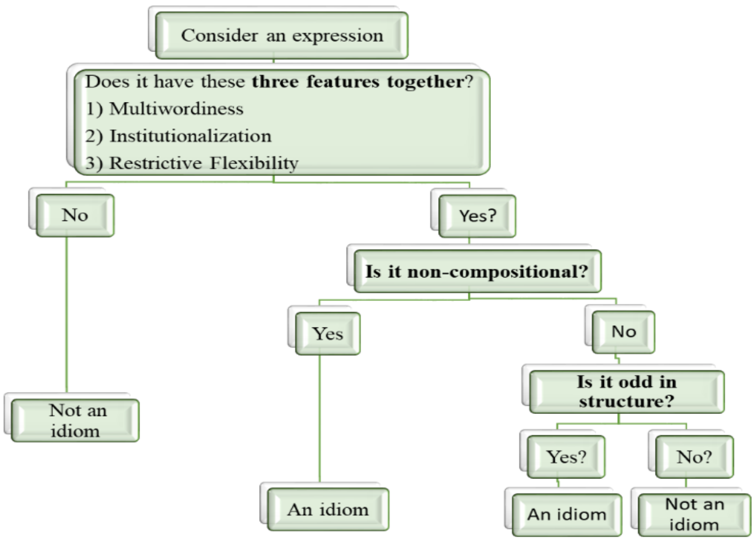
This study investigated idiomatic expressions in applied linguistics RAs using a corpus-based approach, focusing on their formality and usage patterns. A corpus of 640 RAs (5,675,554 tokens) was analyzed to address three research questions: defining academic idioms, determining their frequency, and examining diachronic trends from 1980 to 2020.

Research Question 1: Defining Academic Idioms

To differentiate academic idioms from other multi-word expressions, 240 RAs were manually analyzed, and over 700 expressions were evaluated using dictionaries and expert input. A framework for idiom identification was developed, as shown in Figure 1.

Figure 1

Framework for Idiom Definition and Recognition in Written Academic Corpora



Based on this framework, an expression must meet three initial criteria to be considered an idiom: multi-wordiness (two or more words, less than a sentence), institutionalization (listed in at least one English dictionary), and restrictive flexibility (words cannot be replaced with synonyms without altering meaning). Expressions passing these criteria are further evaluated for non-compositionality (meaning not deducible from components) or odd structure (deviation from grammatical norms). Examples include:

- **Non-compositional idioms**: *hold water*, *then again*, *open the door to* (e.g., "The tendency of linguistics applied will be to dance attendance to whatever tune is currently in theoretical fashion," Edge, 1989, p. 48).

- **Odd in structure idioms**: *for example*, *in turn*, *by and large* (e.g., "By and large, although with some differences, they accepted our interpretation," Alderson & Clapham, 1992, p. 154).
- **Non-idiomatic examples**: *place the blame* (replaceable words), *build on sand* (not institutionalized).

Considering the different nature of idioms in written academic genres such as RAs, in contrast to other more informal genres, this framework can be used for idioms and other similar multiword expressions to be differentiated. Table 1 presents the summary of all the points referred to above in practice by way of examples from the corpus of RAs used in this study.

Table 1

Idiom Identification Test in Academic Genres

Expression	Multi-word	Institutionaliz-ed	Restrictive Flexibility	Non-compositional	Odd Structure	Idiom
Place the blame on sb	Yes	Yes	No (replaceable words)	-		No
Put another brick in the wall	Yes	No	-	-		No
Work way up the academic ladder	Yes	Yes	Yes	No		No
Hold water	Yes	Yes	Yes	Yes		Yes
Then again	Yes	Yes	Yes	Yes		Yes
Give weight to sth	Yes	Yes	No (replaceable words)	-		No
Open a window for sth	Yes	Yes	Yes	No		No
Open the door to sth	Yes	Yes	Yes	Yes		Yes
At any rate	Yes	Yes	Yes	Yes		Yes
For instance	Yes	Yes	Yes	No	Yes	Yes
In the final analysis	Yes	Yes	Yes	Yes		Yes

This framework distinguishes idioms from collocations (e.g., *break free*), metaphors (e.g., *a recipe for failure*), and other expressions by emphasizing their unique properties in academic contexts.

Research Question 2: Frequency of Idioms

Using the developed framework, 567 idioms were identified in 240 RAs and searched across the full corpus using #LancsBox. A total of 42,499 idiom occurrences (564 unique idioms) were found. Table 2 lists idioms with a relative frequency (RF) ≥ 0.10 per 10,000 words, including frequency (f) and range (r, number of texts).

Table 2

Most Frequent Academic Idioms in Applied Linguistics RAs

Idiom	f	rf	r
For example	3565	6.28	407
As well as	2019	3.56	376
In addition	1560	2.75	327
In terms of	1475	2.60	331
In order to	1469	2.59	348
At least	1247	2.20	332
In fact	969	1.71	308
On the other hand	661	1.16	258
In particular	585	1.03	243
For instance	568	1.00	172

Notably, *for example* is idiomatic due to its grammatical deviation (lacking an article before "example"). Non-compositional idioms, such as *shed/shine/throw/cast light on* (f=120, rf=0.21, r=87), were less frequent but significant (Table 3).

Table 3

The Most Frequent Non-Compositional Idioms Used in RAs

Idioms	f	rf	r
Shed/shine/ throw/ cast light on	120	0.21	87
In the light of	60	0.11	45
By/in virtue of	57	0.10	36
In keeping with	41	0.07	33
In the hands of /In sb's hands	36	0.06	32
Come/ bring into play	28	0.05	26
By and large	26	0.05	17
Let alone	24	0.04	20
Come to terms with	20	0.04	11
Bring/come to light	14	0.02	12
Bring sth home	10	0.02	9
Rule(s) of thumb	9	0.02	6
Part and parcel of	8	0.01	7
Be in tune with	8	0.01	8

It stands to reason	8	0.01	8
Come/get to grips with	8	0.01	6
Take sth on board	6	0.01	6
Sweep sth under the carpet	6	0.01	3
Take stock of sth	6	0.01	4
Sit comfortably/uncomfortably/well/with	6	0.01	6
By heart	5	0.009	3
In a nutshell	5	0.009	4
pay lip service to sth	4	0.007	4
Scrape the barrel	4	0.007	1
Shake one's head at sth	3	0.005	3
Stand a chance	3	0.005	3
Leave sb's mark on sth/sb	3	0.005	3
The other way round	3	0.005	3
Fall into the trap of	3	0.005	3
Order of the day	3	0.005	3
Find your voice	3	0.005	3
As much	3	0.005	3
On the right track	3	0.005	3
On a par with	3	0.005	3
Come of age	3	0.005	3

Table 4 lists pedagogically relevant idioms for EAP writing, such as *in terms of* and *in light of*, which are formal and suitable for academic contexts.

Table 4

Pedagogical Academic Idioms

In terms of	In line with	In view of
In light of	Shed/shine/ throw/ cast light on	At best
As it were	Bring/come to light	So to speak
With respect to	In a way (meaning to some extent)	In this regard
In part	So as to	Keep/bear sth in mind
With/in regard to	With reference to	Thus far
As such	In conjunction with	In favour of
As with sb/sth	In this respect	In sb/sth's own right
As yet	In effect	Above all
Bring/come to the fore	In and of itself	For one thing
In a positive/negative light	Sit comfortably/uncomfortably /with	All the same
At any rate	Be in tune with sth	It stands to reason
By way of (example, etc.)	Do justice to sth	Goes without saying…
By virtue of	In reference to	As yet

Give rise to	By and large	When it comes to
That said	By reference to	Then again
By definition	Anything but	For the sake of argument
In keeping with	It is safe to say	Not least
That/this is to say	Stand the test of time	In tandem
Less than	In the final analysis	Not to mention
Bring to bear on/upon	By the same token	What is more
Take issue with sth	Bring sth into focus	

Research Question 3: Diachronic Trends

To examine differences in idiom use from 1980 to 2020, 240 RAs (1,850,559 tokens) were analyzed manually, divided into four decades. Frequencies were normalized to 300,000 words (Table 5).

Table 5

Normalized Frequencies of Idioms per Decade (Per 300,000 Words)

Decade	Articles	Tokens	Idioms	Normed Tokens	Normed Idioms
1980–1989	60	385,194	1259	300,000	980
1990–1999	60	498,665	1882	300,000	1132
2000–2009	60	437,240	1711	300,000	1173
2010–2020	60	531,204	1982	300,000	1119

A Chi-Square test (Table 6) showed no significant difference in idiom use across decades (χ^2 (9) =12.000, p=.213), confirming the null hypothesis.

Table 6

Chi-Square Tests for Idiom Use per Decade

Test	Value	f	Asymp. Sig. (2-sided)
Pearson Chi-Square	12.000		.213
Likelihood Ratio	11.090		.270
Linear-by-Linear Association	1.491		.222

The developed framework effectively identifies academic idioms, revealing their frequent use (42,499 occurrences) in applied linguistics RAs. Idioms like *for example* and *in terms of* are prevalent, while non-compositional idioms like *shed light on* are less common but pedagogically valuable. No significant diachronic changes were observed, suggesting consistent idiom use over time.

Discussion

Academic Idioms Redefined and Recognized

The main purpose of this study was to shed new light on idioms in RAs and develop an operational definition for academic idioms in RAs that have been ignored thus far in the literature. A framework for idiom identification and an operational definition was developed:

> *Idioms are dynamically institutionalized expressions of two words or more, but less than a sentence, whose individual words cannot be replaced with other similar words, and are either non-compositional or odd in structure or both.*

The criteria which were focused on in the framework were: *multi-wordiness* (two words or more but less than a complete sentence), *institutionalization* (i.e., being included in at least one of the dictionaries available in English), *restrictive flexibility (i.e.* The words in the expression cannot be replaced with other synonyms and keep the idiomatic meaning), *non-compositionality* (i.e. the sense of the words in the expression is not included in any of the dictionaries), and/or *oddness in structure* (i.e. The expression deviates from syntactical rules).

It should be borne in mind that an expression must have the first three features together along with either of the last two features, or both of them to be labeled as an idiom.

This definition of academic idioms was attempted to consider the shortcomings of the previous definitions proposed in the literature (e.g., Fernando, 1996; Gramley & Patzold, 2004; Grant & Bauer, 2004). In other words, this definition takes into account a neglected aspect of this linguistic phenomenon which is related to the formal aspect of them. Most of the definitions proposed are related to idioms in informal genres, ignoring that idioms are also used in academic genres (Miller, 2020; Simpson & Mendis, 2003).

In line with the previous studies (e.g., Dobrovol‘skij & Piirainen, 2005; Liontas, 2017; Miller, 2011; Miller, 2020), the size of idioms was opted for to be of at least two words and less than a complete sentence, and ignoring the idiomatic compound words which were included in some studies on idioms (e.g. Katz & Postal, 1964; Makkai, 1972). Moreover, idiomatic phrasal verbs were also excluded since, as Hinkel (2017) argues, “their numbers are so large that their exact or even proximate counts are unknown” (p. 49), in contrast to Liu (2003), who considered phrasal verbs in his study of idioms.

Moreover, *meaning* has been the major criterion on which idioms have been so far defined (e.g., Fernando, 1996; Fraser, 1970; Gramly& Patzold, 2004; Grant & Bauer, 2004; Weinreich, 1969), ignoring the idiosyncrasy of idioms in terms of their syntactic structures. However, the emphasis on this aspect of idioms is in line with Flavel and Flavel (2011) and Gramly and Patzold’s (2004) conceptions of idioms.

More importantly, the definition provided in this study corroborate the findings of the previous corpus-based studies which manifested that non-compositionality as a major feature of idioms should be reconsidered, that non-compositionality is a matter of degree, and some idioms are more non-compositional than others (Geeraert, 2016; Gluckberg et al., 2001; Langlotz, 2006; Moon, 1998; Taylor, 2012; Wulff, 2008; Zyzik, 2011). More to the point, in this definition, non-compositionality is redefined so that it is operational, without involving the subjectivity, as was the case with Grant and Bauer's (2004) framework of idioms. In this framework, the differences between figurative expressions and non-compositional idioms were based on the intuition of a language user.

This study also confirmed that idioms should not be considered totally fixed and frozen structures (Glucksberg et al., 2001; Kyriacou et al., 2020; Langlotz, 2006; Lynn, 2016; Moon, 1998). For example, one of the idioms found in the corpus: *to shed light on something*, was used with many variations by different writers in RAs:

"Qualitative findings ***shed further light on*** *the quantitative results"* (Tedick & Young, 2016, p.20).

"That is, corpus-based research ***sheds new light on*** *some of our most basic assumptions about English"* (Biber et al., 1994, p. 174).

"There was a general agreement that research was necessary into a variety of areas in order ***to throw more light*** *on the nature of language proficiency"* (Alderson & Clapham, 1992, p. 162).

"The diaries ***threw valuable light on*** *aspects of vocabulary teaching which had been covered on the TEFL course"* (Richards, 1992, p.148).

As another example which supports the results of the previous studies in this regard, the idiom *make sense* was also found to be used with variations. For example,

"Use of heavy to modify rain ***makes perfect sense****"* (Liu, 2012, p.18).

"Meanings of strong versus powerful tea ***make very good sense****"* (Liu, 2010, p.18).

Similarly, the idiom: *with reference to* had many variations in its form:

"With specific reference *to L2 grammar teaching it does raise interesting questions about…"* (Borg, 1998, p.31)

"I now turn to consider the description of collocations and idioms in specialized dictionaries, ***with special reference to…"*** (Cowie, 1981, p. 225).

The researcher avoided to use the term *fixedness*; instead, she opted for the term *restrictive flexibility*. This term refers to the fact that words in an idiom cannot be replaced with other similar words and still keeps its idiomatic meaning. This feature is similar to what Gramly and Patzold (2004) refer to as *recurrent semantic contrast.*

As manifested by the examples taken from the corpus, idiomatic collocational patterns and idioms can be differentiated considering these features.

What makes this framework different from the previous frameworks available in the literature is that by using this framework, the neglected group of idioms, and the idioms with oddness in structure, such as *for example* can also be taken into account. This group of idioms has been neglected in previous studies (e.g., Miller, 2020; Simpson & Mendis, 2003).

A List of Academic Idioms

The second research question focused on the frequency of idioms based on the framework and definition developed. A list of the most frequent academic idioms was compiled.

The results indicated that applied linguistics writers of RAs make use of idioms in their RAs. Moreover, the results showed that proficient writers in the field have the confidence to make use of *informal idioms* as well, and this usage is well accepted. In contrast to what is usually stated in EAP and EFL coursebooks, avoiding learners from making use of idioms in written academic genres (Bailey, 2015; Sowton, 2012).

Comparing the list of idioms with the lists of idioms reviewed in the literature shows that there are major differences between them. From the most frequent idioms in academic spoken genres reported by Simpson and Mendis (2003), none of the idioms are in common with the most frequent idioms reported in this study. However, three idioms: *come into play*, *the bottom line*, and *draw a/the line* were common, although with a low frequency.

Regarding the most frequent idioms reported by Grant and Nation (2006), only two idioms: *as well* and *by and large* are in common with the idioms found in this study. Of Chen & Wang's (2016) idioms, only *Achilles' heel* was in common with the idioms in this study.

However, comparing the idioms in Rafatbaksh and Ahmadi (2019) and Rafatbaksh and Ahmadi (2020) with the idioms in this study showed some similarities. Many idioms such as: *in the long run, in the short run, in the light of, as it were, in the final analysis, by and large, pave the way, and in tandem* were in common. This finding can be justified by the point that in these two studies, in contrast to previous studies, a part of the corpus investigated is devoted to academic written genres. Therefore, it stands to reason that there are more similarities between their results and those of this study.

And finally, considering the most frequent idioms with a frequency over 4 pmw in the academic written corpus in Miller (2020) shows that there are many similarities between that list and the results of this study. This can also be explained by the point that one of the corpora used in this study is devoted to academic written genres, although it is not specific to RAs.

The results of the comparisons which were run between the different lists of idioms introduced in the literature and the list in this idioms can be justified by the point that the corpus used in this study is a specialized one, focusing only on one genre and one discipline, in contrast to previous studies which reported generic idiom lists by using different genres in different disciplines (Liu, 2003; Miller, 2020; Rafatbakhsh & Ahamdi, 2020; Simpson & Mendis, 2003).

Given these results, the findings of this study throw into relief the differences that exist between different registers, genres, and disciplines in terms of idiom usage (Alizadeh et al., 2024) and confirm Hyland and Tse (2007) and Durrant's (2009) cautiousness that generic idiom lists will not apply to all disciplines and lexical needs of students vary in different disciplines.

Moreover, the results confirm the results of previous corpus-informed studies making use of huge corpora that although idioms, in general, do not enjoy high frequency compared to the other genres such as fiction, there are idioms more frequent in written academic genres such RAs (Moon, 1998; Biber et al., 1999).

As Biber et al. (1999) state (p.1026), a portion of the idioms found in this study can be called academic idioms specific to academic genres. This finding manifests the importance of dividing idioms into different groups such as *informal idioms*, *formal idioms*, *frequent idioms in spoken academic English*, and *frequent idioms in written academic English*.

The results show that the most frequent idioms found in applied linguistics RAs are relational idioms, which are used as cohesive devices by the authors. It comes as no surprise considering the feature of this academic genre. Most of the idioms in this list might seem totally literal at first sight, but this is not the case with all of them. Some of them are not non-compositional, but they deviate from the grammatical rules of English.

Results show that idioms used in RAs can be divided into three groups: non-compositional idioms, idioms with oddness in structure, and full idioms. Non-compositional idioms are those whose meanings cannot be deduced from the words included. It should be kept in mind that to be included as a non-compositional idiom, an expression must have the three initial features introduced in the previous section.

There are also figurative idioms used by the authors, implying that informal idioms can also be used in this genre. Although these types of idioms have low frequency, as Miller (2020) argues, lack of high frequency is not equal to the lack of significance.

Moreover, comparing the idioms with the lists of idioms reviewed in existing literature, shows that although some idioms are in common in most of these lists, such as *shed light on something*, many of the idioms found in this study are not reported in previous findings. This can be justified by the point that the corpus used in this study is a specialized one, focusing only on one genre, in contrast to previous studies which reported generic idiom lists by using different genres in different disciplines (Liu, 2003; Miller, 2020; Rafatbakhsh & Ahamdi, 2020; Simpson & Mendis, 2003;).

The results of this study throw into relief the differences that exist between different registers, genres, and disciplines in terms of idiom usage and confirm Hyland and Tse (2007) and Durrant's (2009) cautiousness that generic idiom lists will not apply to all disciplines, and that lexical needs of students vary in different disciplines. Moreover, the results confirm the results of previous corpus-informed studies making use of huge corpora that although idioms, in general, do not enjoy high frequency compared to the other genres such as fiction, there are idioms more frequent in written academic genres such RAs (Biber et al., 1999; Moon, 1998).

Any Diachronic Change in the Use of Idioms?

The results which were obtained in response to research question 3 indicate that the use of idioms seems to be a matter of style in writing rather than a matter of time since in each of the four decades, there were some RAs which were richer by far in the use of idioms. An article published in the Journal of Applied Linguistics in 2000 authored by Widdowson (1993) is the most idiomatic of all, including 121 idioms. The next richest RA in terms of idiom usage belongs to decade1, authored by Canale and Swain (1980).

Some writers have the confidence and creativity to make use of idioms in their writings, some do not. This finding might go back to writers' first language. As Baker (2018) argues, "a person's confidence in actively using the idioms and fixed expressions of a foreign language hardly ever matches that of a native speaker" (p. 70). More to the point, as Baker discusses, in some languages, there is a considerable difference between written and spoken registers. In languages such as Arabic and Chinese, language users tend to avoid using idioms in formal discourses since it is assumed that idioms belong to informal registers. Given this point, a writer's tendency to make use of idioms may be influenced by her/his mother tongue. Furthermore, disciplinary variations in considering pragmatic and stylistic aspects of academic writing cannot be neglected. For instance, Khaghaninejad et al. (2022) that objectivity in academic texts varies across academic discilplines.

Conclusion

This study reveals that idioms, often viewed as informal, play a significant role in applied linguistics RAs, a formal academic genre. Though less frequent than other lexical items, idioms are used purposefully and creatively by proficient writers, enhancing cohesion and expressiveness, thus challenging the notion that they are unsuitable for academic writing. A two-stage framework was developed to identify academic idioms (Figure1). Expressions must exhibit multi-wordiness (two or more words, less than a sentence), institutionalization (listed in at least one English dictionary), and restrictive flexibility (non-replaceable words). Qualifying expressions are then evaluated for non-compositionality (meaning not deducible from components) or odd structure (deviation from grammatical norms). This framework, exemplified in Table1, distinguishes idioms from collocations and metaphors.

The operational definition of idioms as dynamically institutionalized expressions of two or more words, non-replaceable, and either non-compositional or odd in structure, provides clarity for corpus-based studies. Analysis of a 640-RA corpus (5,675,554 tokens) identified 42,499 idiom occurrences (564 unique idioms), with relational idioms like *for example* (f=3565, rf=6.28) and *in terms of* (f=1475, rf=2.60) serving as cohesive devices. Informal idioms (e.g., *Achilles heel*) were less frequent but accepted in prestigious journals, reflecting authorial style and idiomatic competence. No significant diachronic differences in idiom frequency were found across 1980–2020 (χ^2 (9) =12.000, p=.213), suggesting usage depends on writers' stylistic preferences rather than temporal trends. The framework addresses a literature gap by distinguishing idioms from other multi-word expressions, enhancing understanding of their role in formal genres. For EFL/EAP contexts, it aids teachers in clarifying idioms versus collocations or proverbs, improving learners' idiomatic competence (Liontas, 2017). The idiom list, particularly formal ones like *in light of*, can be integrated into EAP writing courses and textbooks to enhance authenticity and reduce cognitive load (Adel & Erman, 2012). Authentic RA examples (e.g., Widdowson, 1993) model creative idiom use, and the list supports translation courses addressing idiom challenges. Limitations include the manual analysis of only 240 RAs due to time constraints and the lack of comprehensive inter-rater reliability due to idiom definition complexities. Future research could explore idioms in other disciplines, compare native and non-native writers, investigate idioms in other languages (e.g., Farsi), or examine novice versus proficient writers. Research on other academic genres (e.g., theses, editorials) and idiom functions (e.g., evaluation, emphasis) is also recommended.

References

Adel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes*, *31*(2), 81-92. https://doi.org/10.1016/j.esp.2011.08.004

Alizadeh, A., Kuhi, D., & Hadidi, Y. (2024). Structural and functional differences of lexical bundles between Hard Science and Soft Science researchers: A comparative corpus-driven study. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, *12*(2), 87-114. https://doi.org/10.22049/jalda.2024.29174.1630

Bailey, S. (2015). *Academic writing: A handbook for international students* (4th ed.). Routledge.

Baker, M. (2018). *In other words: A coursebook on translation*. Routledge.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Longman.

Boers, F., Eyckmans, J., & Stengers, H. (2007). Presenting figurative idioms with a touch of etymology: More than mere mnemonics?. *Language Teaching Research*, *11*(1), 43-62. https://doi.org/10.1177%2F1362168806072460

Brenner, G. (2011). *Webster's new world American idioms handbook*. Mifflin Harcourt.

Brezina, V., Timperley, M., & McEnery, A. (2018). #LancsBox v. 4.x. [software package].

Brezina, V., Timperley, M., & McEnery, A. (2020). #LancsBox v. 4.x. [software package]

Chen, Y., & Wang, L. (2016). Opening Pandora's box: A corpus-based study of idioms in ELT materials. *The Journal of Asia TEFL*, *13*(4), 381-39. http://dx.doi.org/10.18823/asiatefl.2016.13.4.10.381

Cooper, T. C. (1998). Teaching idioms. *Foreign Language Annals*, *31*(2), 255-266. https://doi.org/10.1111/j.1944-9720.1998.tb00572.x

Dobrovol'skij, D & Piirainen, E. (2005). *Figurative language: Cross-cultural and cross-linguistic perspectives*. Elsevier Ltd.

Durrant, P. (2009). Investigating the viability of a collocation list for students of English for academic purposes. *English for Specific Purposes*, *28*, 157-169. https://doi.org/10.1016/j.esp.2009.02.002

Fernando, C. (1996). *Idioms and idiomaticity*. Oxford University Press.

Fraser, B. (1970). Idioms within a transformational grammar. *Foundations of Language*, *6*(1), 22-42.

Geeraert, K. D. (2016). *Climbing on the bandwagon of idiomatic variation: A multi-methodological approach*. [Doctoral dissertation, University of Alberta]. University of Alberta Libraries. https://doi.org/10.7939/R3222RB4F

Glucksberg, S., McGlone, M. S., Grodzinsky, Y., & Amunts, K. (2001). *Understanding figurative language: From metaphor to idioms* (No. 36). Oxford University Press.

Gramley, S., & Pátzold, M. (2004). *A survey of modern English*. Routledge.

Grant, L., & Bauer, L. (2004). Criteria for re-defining idioms: Are we barking up the wrong tree? *Applied Linguistics*, *25* (1), 38-61. https://doi.org/10.1093/applin/25.1.38

Grant, L., & Nation, I.S. P. (2006). How many idioms are there in English? *ITL-International Journal of Applied Linguistics, 151*, 1-14. https://doi.org/10.2143/ITL.151.0.2015219

Hinkel, E. (2017). Teaching idiomatic expressions and phrases: Insights and techniques. *Iranian Journal of Language Teaching Research*, *5*(3), 45-59. http://dx.doi.org/10.30466/ijltr.2017.20304

Hocket. C. F. (1958). *A Course in modern linguistics*. Macmillan.

Hyland, K. (2009). *Academic discourse: English in a global context*. A&C Black.

Hyland, K. L., & Paltridge, B. (Eds.). (2011). *Continuum companion to discourse*. Continuum.

Hyland, K., & Tse, P. (2007). Is there an "academic vocabulary"? *TESOL Quarterly*, *41*(2), 235-253. https://doi.org/10.1002/j.1545-7249.2007.tb00058.x

Katz, J. J., & Postal, P. M. (1964). *An integrated theory of linguistic descriptions*. M.I.T. Press.

Khaghaninejad, M. S., Azarian, M., & Javanmardi, F. (2022). A corpus-based cross-disciplinary analysis of objectivity manifestations in academic texts. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, *10*(1), 119-140. https://dx.doi.org/10.22049/jalda.2022.27325.1333

Khonbi, Z., & Sadeghi, K. (2017). Improving English language learners' idiomatic competence: does mode of teaching play a role? *Iranian Journal of Language Teaching Research*, *5*(3), 61-79. https://dx.doi.org/10.30466/ijltr.2017.20305

Kyriacou, M., Conklin, K., & Thompson, D. (2020). Passivizability of idioms: Has the wrong tree been barked up?. *Language and Speech*, *63*(2), 404-435. https://doi.org/10.1177%2F0023830919847691

Langlotz, A. (2006). *Idiomatic creativity: A cognitive-linguistic model of idiom-representation and idiom-variation in English*. John Benjamins Publishing Company.

Leah, C. (2014). *Idioms-grammaticality and figurativeness*. http://www.theroundtable.ro/Current/Language/Claudia_Leah_Idioms_Grammaticality_and_Figurativeness.pdf

Liontas, J. I. (2017). Why teach idioms? A Challenge to the profession. *Iranian Journal of Language Teaching Research*, *5*(3), 5-25. https://dx.doi.org/10.30466/ijltr.2017.20302

Liu, D. (2003). The most frequently used spoken American English idioms: A corpus analysis and its implications. *TESOL Quarterly, 38*(4), 671-700. https://doi.org/10.2307/3588217

Liu, D. (2012). The most frequently-used multi word constructions in academic written English: A multi-corpus study. *English for Specific purposes*, *31*, 25-35. https://doi.org/10.1016/j.esp.2011.07.002

Lynn, E. M. (2016). *Getting all the ducks in a row: Towards a method for the consolidation of English idioms.* [Master's thesis, Brigham Young University]. BYU Scholars Archive. https://scholarsarchive.byu.edu/etd/6014/

Makkai, A. (1972). *Idiom structure in English*. De Gruyter Mouton.

McCarthy, M. J. (1998). *Spoken language and applied linguistics*. Cambridge University Press.

McCarthy, M. J. & O'Dell, F. (2010) *English idioms in use*. Cambridge University Press.

McCarthy, M. J. & O'Dell, F. (2008). *Academic vocabulary in use*. Cambridge University Press.

Miller, J. (2011). *Never judge a wolf by its cover*: *An investigation into the relevance of phrasemes included in advanced learners' dictionaries for learners of English as an additional language in Australia.* [Unpublished doctoral dissertation]. Flinders University

Miller, J. (2020). The bottom line: Are idioms used in English academic speech and writing? *Journal of English for Academic Purposes*, *43*, 1-14. https://doi.org/10.1016/j.jeap.2019.100810

Moon, R. (1998). *Fixed expressions and idioms in English.* Clarendon Press.

Nunberg, G. D. (1978). *The pragmatics of reference*. Indiana University Linguistics Club.

O'Dell, F., & McCarthy, M. (2017). *English collocations in use: Advanced, how words work together for fluent and natural English; self-study and classroom use.* Cambridge University Press.

O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching.* Cambridge University Press.

Panou, D. (2014). *Idiom translation in the financial press: a corpus-based study*. Cambridge Scholars Publishing.

Rafatbakhsh, E., & Ahmadi, A. (2019). A thematic corpus-based study of idioms in the Corpus of Contemporary American English. *Asian-Pacific Journal of Second and Foreign Language Education*, *4*(1), 4-11. https://doi.org/10.1186/s40862-019-0076-4

Rafatbakhsh, E., & Ahmadi, A. (2020). The most frequent idioms used in contemporary American English: A corpus-based study. *Applied Research on English Language*, *9*(2), 205-228. https://dx.doi.org/10.22108/are.2019.114449.1389

Richards, J. C., & Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics.* Longman.

Simpson, R., & Mendis, D. (2003). A corpus-based study of idioms in academic speech. *TESOL Quarterly*, *37*(3), 419-441. https://doi.org/10.2307/3588398

Sinclair, J. (1991). *Corpus, concordance, and collocation*. Oxford University Press.

Simpson, A.J., & Weiner, E. S. C. (1989). *The Oxford English dictionary*. Oxford University Press.

Sowton, C. (2012). *50 steps to improving your academic writing*. Garnet Education.

Taylor, J. R. (2012). *Mental corpus: How language is represented in the mind*. Oxford University Press.

Van Enk, A., & Power, K. (2017). What is a research article? Genre variability and data selection in genre research. *Journal of English for Academic Purposes*, *29*, 1-11. https://doi.org/10.1016/j.jeap.2017.07.002

Weinreich, U. (1969). Problems in the analysis of idioms. In J. Puhvel (Ed.), *The substance and structure of language* (pp. 23-81). University of California Press.

Wulff, S. (2008). *Rethinking idiomaticity: A usage-based approach*. A&C Black.

Zyzik, E. (2011). Second language idiom learning: The effects of lexical knowledge and Pedagogical sequencing. *Language Teaching Research*, *15*(4), 413-433. https://doi.org/10.1177%2F1362168811412025

Corpus References

Alderson, J. C., & Clapham, C. (1992). Applied linguistics and language testing: A case study of the ELTS test. *Applied Linguistics*, *13*(2), 149-167. https://doi.org/10.1093/applin/13.2.149

Biber, D., Conrad, S., & Reppen, R. (1994). Corpus-based approaches to issues in applied linguistics. *Applied Linguistics*, *15*(2), 169-189. https://doi.org/10.1093/applin/15.2.169

Borg, S. (1998). Teachers' pedagogical systems and grammar teaching: A qualitative study. *TESOL Quarterly*, *32*(1), 9-38. https://doi.org/10.2307/3587900

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, *1*(1), 1-47. https://doi.org/10.1093/applin/I.1.1

Cowie, A. P. (1981). The treatment of collocations and idioms in learners' dictionaries. *Applied Linguistics*, *2*(3), 223-235. https://doi.org/10.1093/applin/II.3.223

Edge, J. (1989). Ablocutionary value: On the application of language teaching to linguistics. *Applied Linguistics*, *10*(4), 407-417. https://doi.org/10.1093/applin/10.4.407

Liu, D. (2010). Going beyond patterns: Involving cognitive analysis in the learning of collocations. *TESOL Quarterly*, *44*(1), 4-30. https://doi.org/10.5054/tq.2010.214046

Richards, K. (1992). Pepys into a TEFL course. *ELT Journal*, *46*(2), 144-152. https://doi.org/10.1093/elt/46.2.144

Tedick, D. J., & Young, A. I. (2016). Fifth grade two-way immersion students' responses to form-focused instruction. *Applied Linguistics*, *37*(6), 784-807. https://doi.org/10.1093/applin/amu066

Widdowson, H. G. (1993). Proper words in proper places. *ELT Journal*, *47*(4), 317-329. https://doi.org/10.1093/elt/47.4.317

Authors' Biography

Elham Farahani received her BA in English Language and Literature from the University of Qom, and her MA and PhD degrees in TEFL from Arak University. She has served as a faculty member from 2018 to 2022 in the Department of English Language at Mofid University. In 2023, she joined Arak University, where she is currently an Assistant Professor. Her areas of interest include Corpus Linguistics, English for Academic Purposes, English for Research Publication Purposes, and (Critical) Discourse Analysis.

Hooshang Yazdani, Associate Professor of TEFL at Arak University, studied TEFL as an undergraduate at Shiraz University and went on to do an MA in the same field. He continued the same field of study for PhD. He conducted a piece of research on a new way of evaluating EFL readers' reading comprehension abilities for his MA Degree and investigated the effects of different variables on reading strategy choice and comprehension for PhD Degree. He has taught different basic courses related to language, ESP and Applied Linguistics for about five years before starting his PhD abroad in 1993 and nineteen years since holding the degree in 2000. His research interests mainly lie in Discourse Analysis, Vocabulary Acquisition, Individual Differences & SLA.

Volume 14, Issue 1
Winter and Spring, 2026
pp. 99-122

Journal of Applied Linguistics and Applied Literature: Dynamics and Advances

An Exploration into Iraqi EFL Learners' Acceptance of ChatGPT for Language Learning: A Technology Acceptance Model Study

Ahmed Rawdhan Salman[1], Saman Ebadi[2*], Sajad Velayati[3]

[1]*Postdoctoral Researcher in ELT, Razi University, Kermanshah, Iran*
ahmed.rawdhan@uomus.edu.iq

[2]*Corresponding author: Professor of TEFL, Department of English Language and Linguistics, Faculty of Literature and Humanities, Razi University, Kermanshah, Iran; ORCID: 0000-0001-8623-7751*
samanebadi@gmail.com

[3]*Phd Candidate in ELT, Razi University, Kermanshah, Iran*
velayati.sajad@gmail.com

Abstract

This study adopted a sequential explanatory mixed-methods approach to investigate the variables that govern the acceptance and use of ChatGPT for learning English by 220 Iraqi university students. According to the Technology Acceptance Model (TAM), the study used a survey to measure perceived usefulness, ease of use, attitude, behavioral intention and actual use, as well as perceived enjoyment, facilitating conditions and technological complexity. Later, interviews were conducted in a semi-structured manner. The quantitative analysis revealed that while gender was not a significant predictor of attitude, it influenced actual usage rates. More significantly, prior experience with technology emerged as a critical dividing line, and students with high digital readiness reported significantly more positive perceptions across all constructs, particularly regarding facilitating conditions. Qualitative data revealed a clear pattern showing that English as a Foreign Language (EFL) learners see ChatGPT as a tool that enhances proficiency, increases motivation, and provides personalized learning suggestions. Although there are some benefits of self-assessment, there are nonetheless substantial concerns surrounding the absence of dynamic interactional feedback and the danger of developing over-reliance and ethical considerations. The findings of the study showed that Iraqi EFL learners viewed the ChatGPT as a beneficial tool. However, to ensure its effective and fair implementation, pedagogical frameworks must be designed to scaffold digital literacy and provide clear guidelines to control its practical and ethical shortcomings.

Keywords: ChatGPT, gender, Iraqi EFL learners, mixed-methods approach, Technology Acceptance Model (TAM)

ARTICLE INFO
Research Article
Received: Sunday, November 16, 2025
Accepted: Sunday, February 15, 2026
Published: Wedensday, April 1, 2026
Available Online: Sunday, February 15, 2026
DOI: https://doi.org/10.22049/jalda.2026.31224.1923
Online ISSN: 2821-0204; Print ISSN: 28208986

Introduction

The rise of generative artificial intelligence (AI), which has drastically changed various sectors (Alshahrani et al., 2022; Dennehy et al., 2022; Dwivedi et al., 2021), has affected global education. Since November 2022, there has been serious debate on whether ChatGPT by Open AI can be used in the teaching of EFL or whether it can disrupt language learning (Li et al., 2023; Shiri, 2023; Zhuo et al., 2023). In contexts where EFL writers do not often have access to authentic use of language, the capabilities of instant feedback generation and personalized text and writing support provided by ChatGPT can be beneficial to EFL contexts (Bin-Hady et al., 2025; Hong, 2023; Kostka & Toncelli, 2023; Song & Song, 2023). However, its adoption remains contentious. According to Kohnke et al. (2023) and Liu et al. (2024), it might help enhance engagement and task accomplishment. The use of AI tools may create a risk to academic integrity and ethical issues because of their overreliance (Abdaljaleel et al., 2024; Sallam, 2023). The utilization of ChatGPT by students for educational purposes is increasing; however, there is a lack of clarity regarding whether it is being employed effectively. Therefore, it is imperative to further investigate this issue.

Reports have highlighted that the ChatGPT can be effective in improving a student's language, especially writing and vocabulary acquisition. AI helps them with personalized practice exercises and explanatory feedback on their language use (Han et al., 2023; Javaid et al., 2023; Kasneci et al., 2023; Koraishi, 2023; Xiao & Zhi, 2023). An exploratory study revealed that students found ChatGPT to be a good learning companion. Furthermore, students modify prompts for better learning (Xiao & Zhi, 2023). With 24/7 availability, ChatGPT helps students enhance their language ability because of critical feedback that stimulates curiosity and proper input difficulty (Kohnke et al., 2023). Studies suggest that it helps learners engage better by self-editing while producing formal texts (Özçelik & Ekşi, 2024). Jeon et al. (2023) reviewed the literature and found that large language models may be more capable than just performing simple conversational tasks, and further research should investigate how to effectively combine visual, text, and auditory modality for language learning.

This study examines the adoption of ChatGPT using the Technology Acceptance Model (TAM) as its framework. According to Chuttur (2009), TAM is probably the most widely known model that explains how technology acceptance takes place. Furthermore, it includes perceived usefulness (the degree to which a user believes that using a technology will enhance his/her performance) as well as perceived ease of use (the degree to which using the technology will be free of effort). Al-Nuaimi and Al-Emran (2021) also reported these two factors as being the most used determinants of technology acceptance. TAM was developed by Davis (1986, 1989). Recent studies have applied TAM with hedonic attributes and contextual factors such as perceived enjoyment (intrinsic) and facilitating conditions (institutional support) relevant in AI-driven learning spaces (Elshaer et al., 2024; Venkatesh et al., 2003). Liu et al. (2024) assert that the enjoyment of the chatbot is a significant factor influencing EFL learners' adoption of ChatGPT. Nonetheless, the primary obstacle is the technological complexity involved.

The Research Gap: AI Adoption in the Iraqi Educational Context

While research on ChatGPT acceptance in EFL is increasing globally, it remains uneven. A growing body of literature exists for East Asia and the Gulf region (e.g., Saudi Arabia and Kuwait) regarding gender and digital literacy in AI adoption. However, Iraq represents a distinct context characterized by unique infrastructural challenges, including intermittent electricity and variable internet connectivity, alongside a higher education system currently undergoing post-conflict reconstruction and digitization. Several authors made a call for culturally responsive studies that take into account variables like gender, prior experience with technology, and local facilitating conditions when modelling acceptance (Alotaibi et al., 2025; Ameen et al., 2021; Bodani et al., 2023; Zhang et al., 2023; Sallam, 2023). In addition, research has highlighted the need to tackle ethical concerns, accuracy, and responsible integration for pedagogically sound use in EFL learning (Cotton et al., 2024; Mhlanga, 2023; Rahman & Watanobe, 2023; Farr, 2024; Liu et al., 2024).

Research in East Asia and the Middle East has considered elements such as gender and digital literacy (Alotaibi et al., 2025; Hwang et al., 2025), but gaps remain. Gender differences in technology adoption have been extensively studied. For example, Elshaer et al. (2024) reported higher self-efficacy of male learners with the use of AI tools. However, other studies showed no gender effect (Panagiotarou et al., 2020). Having previous experience with technology can help lower barriers to adoption. In Iraq, where English proficiency is crucial for academic and professional success, little is known about these dynamics. There are virtually no empirical studies on the perception of Iraqi EFL learners regarding ChatGPT as a teaching tool. This gap is critical because theories of technology acceptance validated in digitally mature environments may not fully explain adoption behaviors in contexts where Facilitating Conditions are precarious. Furthermore, in Iraq, where English proficiency is increasingly viewed as a gateway for global economic participation, understanding how students leverage AI to bridge educational gaps is essential. To address these gaps, this study applies an extended TAM framework to assess ChatGPT acceptance among Iraqi EFL learners in order to direct the ethical and effective integration of ChatGPT in Iraqi EFL contexts. This study was guided by the following research questions:

1. What are Iraqi EFL students' perceptions and experiences of ChatGPT across TAM dimensions, including perceived ease of use, perceived usefulness, attitudes, behavioral intention, actual usage, perceived enjoyment, facilitating conditions, and technological complexity?
2. Are there differences in participants' perceptions of the ChatGPT across gender groups (90 males vs. 130 females)?
3. Are there significant differences in the perceptions of users with and without prior experience with language-learning technology?
4. How do Iraqi EFL learners perceive the advantages and difficulties they encounter when utilizing ChatGPT for language learning?

Literature Review

ChatGPT in EFL Learning: Global and Regional Perspectives

The use of tools of generative AI, such as ChatGPT, has rapidly transformed English as a Foreign Language education at a fast rate (Xiao & Zhi, 2023). For instance, they provide personalized assistance, instant feedback, and opportunities for autonomous practice. Research has shown positive effects on students' writing and vocabulary, as well as custom explanations and practice (Han et al., 2023; Javaid et al., 2023; Kasneci et al., 2023; Koraishi, 2023). Students often use ChatGPT as a friend, where they give strategic prompts to obtain the best learning outcomes (Xiao & Zhi, 2023). Meanwhile, classroom-specific work points to effective feedback that promotes engagement and progress (Kohnke et al., 2023). According to case-based and mixed-methods evidence, there are also observed gains in self-editing and participation, especially for formal writing tasks (Punar Özçelik & Yangın Ekşi, 2024). In EFL contexts that have not provided enough authentic input, ChatGPT can be continuously available to extend practice opportunities and scaffold grammar, vocabulary, and academic writing (Bin-Hady et. al, 2025; Hong, 2023; Kostka & Toncelli, 2023). Previous studies have also highlighted the growing capacities of large language models and the need to understand multimodal integration in learning (Jeon et al., 2023).

Simultaneously, excessive reliance on AI can undermine creativity and thinking. According to Ali et al. (2023), ChatGPT should be used as an additional tool, not as a pedagogical replacement. There is a growing concern about ethical issues relating to academic integrity (Cotton et al., 2024; Mhlanga, 2023; Rahman & Watanobe, 2023; Utebayeva, 2024). Farr (2024) and Liu et al. (2024) mention other challenges, such as the inability to contextualize and biased and inaccurate answers. Further studies warn that automated feedback can produce easy and sometimes misleading responses that must be moderated (Kim et al., 2023; Weidener & Fischer, 2023). Furthermore, studies have indicated that responses can be incorrect or too simplistic (Kotsis, 2024; Xiong, 2024). Across the Middle East, research highlights how perceptions of usefulness, local digital literacy, and social attitudes shape adoption (Alotaibi et al., 2025; Sallam, 2023). These trends have motivated local studies on ChatGPT and its role in education.

Technology Acceptance Model (TAM) and Its Application to AI Tools

The TAM, developed by Davis (1989), posits that perceived usefulness (PU) and perceived ease of use (PEU) shape users' attitudes, intentions, and technology adoption. Recent studies extend TAM to AI-driven tools, incorporating factors such as perceived enjoyment and facilitating conditions to better predict adoption behaviors (Elshaer et al., 2024; Venkatesh et al., 2003). Studies show that students' intentions to adopt e-learning platforms are affected by perceived ease and usefulness (Granić & Marangunić, 2019). TAM extensions include social influence and facilitating conditions, which positively affect behavioral intentions (Khanchel, 2023). Perceived enjoyment influences ChatGPT adoption among EFL learners, whereas technological complexity hinders uptake (Liu et al., 2024). Hwang et al. (2025) demonstrate that output quality enhances perceived usefulness in AI-driven learning environments. Recent research using TAM found that Norwegian students' intention to use the ChatGPT was influenced by behavioral intention and performance expectancy (Grassini et al., 2024). However, the impact of cultural and geographic factors on technology adoption remains understudied, particularly regarding AI tool acceptance among Kuwaiti EFL learners (Ma et al., 2024). Studies have indicated that gender, context, and experience are important factors in AI adoption (Ameen et al., 2021; Bodani et al., 2023; Zhang et al., 2023).

Sociocultural and Individual Influences on ChatGPT Adoption

Contextual elements and personal differences influence the choice to implement AI tools. Factors, such as digital literacy, cultural perspectives, and infrastructure, have a significant impact on these areas. This is supported by regional studies conducted in East Asia and the Middle East (Alotaibi et al. 2025; Hwang et al. 2025; Sallam, 2023). Studies on gender deliver mixed results. For instance, some suggest that males have higher AI self-efficacy or more positive attitudes toward technology, possibly due to exposure. However, other studies have shown non-significant effects of gender (Draxler et al., 2023; Elshaer et al., 2024; Sindermann et al., 2022). Prior technology experience is a consistent predictor of positive perceptions of usefulness and ease of use that positively affects acceptance and continued use (Panagiotarou et al., 2020; Liu et al., 2024). According to Wang et al. (2023), knowledge and attitudes influence emotional responses toward AI, which means that emotions do not lead to acceptance by themselves. Research suggests that social influence and facilitating conditions can moderate behavioral intentions (Ma et al., 2024; Porto et al., 2019; Schoonenboom, 2012; Tandon et al., 2019; Tarhini et al., 2015) in conjunction with locality and cultural norms. Thus, it is important to study AI acceptance in national and educational ecologies.

Methodology

Research Design

Following Creswell & Clark (2007), this study adopted an explanatory sequential mixed-methods design to examine Iraqi EFL learners' acceptance of the ChatGPT. This approach was implemented in two distinct phases. The first phase involved collecting quantitative data through an online survey. The analysis identified general trends and statistical patterns. The second phase involved qualitative data collection through semi-structured interviews to explain and elaborately deal with quantitative data. This approach allows for a "thick description" of the statistical results, particularly regarding the reasons behind observed gender and experience differences. The use of both types of data enhanced the validity and reliability of the study design.

Participants and Context

The study was conducted in an Iraqi higher education setting, specifically EFL undergraduate students. Convenience and snowball sampling methods were used to recruit participants from various universities across Iraq. This non-probability sampling strategy was deemed appropriate given the exploratory nature of the study and the lack of a centralized national registry of students using AI tools. Furthermore, in a post-conflict context where institutional data are often fragmented, snowball sampling allows access to a broader range of students through peer networks.

A total of 220 students taking part in the quantitative survey phase involved 90 male and 130 female EFL learners. The decision to survey EFL majors specifically ensures that the participants possess the necessary English proficiency (approximating CEFR B1-B2) to comprehend the survey items and the AI tool itself, as the Iraqi curriculum for English majors includes extensive coursework in linguistics, literature, and compositions taught entirely in English. A purposive sampling method was employed in the initial quantitative phase. This sampling strategy aimed to ensure equal gender representation and a comprehensive range of attitudes, both positive and negative, as well as the usage patterns identified in the survey. The researchers adhered to ethical standards in their work. Before data collection, the participants were fully informed of the study's purpose, procedures, and their rights. The researchers assured the study participants that they had the right to withdraw from the study at any time without penalty. Furthermore, it was assured that the information collected from them would be anonymized. All participants provided written informed consent before the commencement of the study.

Data Collection Instruments and Procedure

Phase 1: Quantitative Survey

This study employed a dual-phase approach to data collection, incorporating both a quantitative online survey and qualitative semi-structured interviews. The quantitative data were gathered through an online survey (Appendix A) created using Google Forms, which began with demographic questions using a nominal scale to collect information on participants' gender, age, field of study, English proficiency, academic degree, frequency of ChatGPT usage, and previous experience with language learning technology. The second section utilized a five-point Likert scale, ranging from 'strongly disagree' (1) to 'strongly agree' (5), to assess eight TAM constructs adapted from existing TAM literature (Alotumi, 2022; Bai et al., 2022; Liu et al., 2023). These constructs evaluated participants' perceived ease of use (PEU, five items), PU (five items), attitude (A, five items), behavioral/continuance intention (BI, five items), actual use (AU, five items), PE (three items), facilitating conditions (FC, eight items), and TC (five items) of the ChatGPT for learning English. The adaptation of survey items from previous studies is justified by the robustness of TAM's key elements, which are applicable across various contexts (Alharbi & Drew, 2014) and among diverse users and technologies (Nadri et al., 2018; Solano-Lorente et al., 2013). Participation was voluntary, and informed consent was obtained by presenting participants with an introductory statement at the beginning of the survey, which explained the study's purpose, their right to participate in the survey, or follow-up interviews, and assured their confidentiality (anonymous information sharing in the study).

Reliability and Validity of the Questionnaire

To ensure the psychometric rigor of the adapted instrument, we assessed internal consistency using Cronbach’s alpha), Composite Reliability (CR), and Average Variance Extracted (AVE). As shown in Table 1, all constructs demonstrated robust reliability ($\alpha > 0.70$) and convergent validity (AVE > 0.50), satisfying the criteria for educational technology research.

Table 1

Psychometric Properties of the Measurement Scales

Construct	Items	Cronbach's α	CR	AVE
Perceived Usefulness (PU)	5	0.88	0.91	0.68
Perceived Ease of Use (PEU)	5	0.86	0.89	0.62
Attitude (ATT)	5	0.84	0.88	0.60
Behavioral Intention (BI)	5	0.87	0.90	0.65
Actual Use (AU)	5	0.81	0.85	0.54
Perceived Enjoyment (PE)	3	0.82	0.86	0.67
Facilitating Conditions (FC)	8	0.79	0.83	0.51
Technological Complexity (TC)	5	0.78	0.82	0.53

Phase 2: Qualitative Interviews

A group of 20 students who volunteered for the additional study provided their email addresses at the conclusion of the survey. Purposive sampling was employed to select participants with diverse academic backgrounds and varying levels of ChatGPT usage. (see Table 2). Semi-structured interviews were conducted using WhatsApp voice chat to ensure ease and accessibility for the participants. Each interview, lasting between 15 minutes, was recorded with participants' consent for later analysis. While English was the primary language, participants were encouraged to code-switch to Arabic when necessary to express complex ethical or emotional concepts, ensuring the validity of the qualitative data. Informed consent was obtained before each interview. The semi-structured interview guide (Appendix B) explored participants' experiences with ChatGPT in learning English, focusing on their assessments of its benefits and drawbacks, factors affecting its use, and any challenges they faced.

Results

Quantitative Data Analysis

Addressing the First Research Question

The initial analysis focused on describing the central tendencies and variability of students' perceptions across the core Technology Acceptance Model (TAM) components. Table 1 displays the descriptive statistics for the TAM components in descending order.

Table 1

Descriptive Statistics for TAM Components in a Descending Order

TAM Components	N	Min	Max	Mean	SD.	Var
PU	220	1.00	5.00	3.77	.88	.79
PEU	220	1.00	5.00	3.73	.95	.90
Attitude	220	1.00	5.00	3.67	.92	.85
BI	220	1.00	5.00	3.64	.92	.85
PE	220	1.00	5.00	3.61	1.00	1.00
AU	220	1.00	5.00	3.58	1.01	1.04
FC	220	1.00	5.00	3.38	.87	.76
TC	220	1.00	5.00	2.76	1.09	1.19

As indicated in Table 1, Perceived Usefulness (PU) obtained the highest mean score (M = 3.77, SD = 0.88), followed by Perceived Ease of Use (PEU) (M = 3.73, SD = 0.95). Attitudes (M = 3.67, SD = 0.92), Behavioral Intention (BI) (M = 3.64, SD = 0.92), Perceived Enjoyment (PE) (M = 3.61, SD = 1.00), and Actual Usage (AU) (M = 3.58, SD = 1.01) clustered around the midpoint of the scale. Facilitating Conditions (FC) scored lower (M = 3.38, SD = 0.87), while Technological Complexity (TC) was perceived as the most important barrier, which yielded the lowest mean score (M = 2.76, SD = 1.09). To complement the tabular data, a visual representation (Figure 1) was created to illustrate the hierarchical order of the mean scores for each TAM component.

Figure 1

TAM Components Means

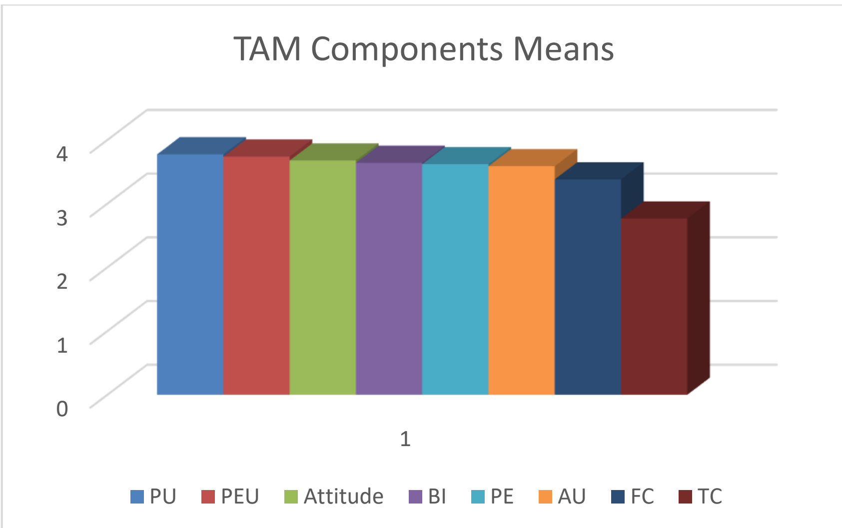


Figure 1 shows the mean values presented in Table 1. The descending bar chart effectively visualizes the continuum of student perceptions from the highest mean (PU) to the lowest (TC).

Addressing the Second Research Question

Prior to conducting inferential tests to compare the groups, it was essential to examine the descriptive statistics and test the assumption of normality for each TAM component separately for male and female participants. Table 2 portrays the descriptive statistics and normality test results for TAM components and TAM total across gender.

Table 2

Descriptive Statistics and Normality Test Results for TAM Components and TAM Total across Gender

TAM Components	Gender	No	Mean	SD	Var	Kolmogorov-Smirnova			Shapiro-Wilk		
						Statistic	df	Sig.	Statistic	df	Sig.
PEU Mean	Male	90	3.74	1.00	1.01	.14	90	.00	.91	90	.00
	Female	130	3.73	.91	0.83	.11	130	.00	.94	130	.00
PU Mean	Male	90	3.83	.91	0.84	.13	90	.00	.90	90	.00
	Female	130	3.73	.87	0.75	.09	130	.00	.95	130	.00
Attitude Mean	Male	90	3.72	.99	0.98	.11	90	.00	.93	90	.00
	Female	130	3.64	.87	0.76	.10	130	.00	.95	130	.00
BI Mean	Male	90	3.73	.95	0.90	.12	90	.00	.92	90	.00
	Female	130	3.57	.90	0.81	.09	130	.01	.96	130	.00
AU Mean	Male	90	3.73	.99	0.98	.14	90	.00	.89	90	.00
	Female	130	3.48	1.03	1.06	.11	130	.00	.94	130	.00
PE Mean	Male	90	3.70	.97	0.94	.15	90	.00	.93	90	.00
	Female	130	3.55	1.01	1.03	.15	130	.00	.93	130	.00
FC Mean	Male	90	3.50	.86	0.75	.12	90	.00	.96	90	.02
	Female	130	3.29	.87	0.77	.11	130	.00	.97	130	.01
TC Mean	Male	90	2.81	1.08	1.17	.12	90	.00	.95	90	.00
	Female	130	2.73	1.10	1.21	.08	130	.01	.95	130	.00
TAM Total	Male	90	3.60	.74	.56	.200	90	.000	.905	90	.000
	Female	130	3.47	.68	.47	.095	130	.006	.972	130	.009

As displayed in Table 2, male participants reported slightly higher mean scores than female participants across all TAM components and for the TAM Total score. The most notable differences appeared for Actual Usage (AU) (Males: M = 3.73; Females: M = 3.48) and Facilitating Conditions (FC) (Males: M = 3.50; Females: M = 3.29). However, the normality tests for both groups across all variables yielded statistically significant results ($p < .05$ for most tests, except a few $p > .05$), which leads to the violation of normality for all the comparison pairs. Accordingly, the Mann-Whitney U test was employed for subsequent group comparisons, the results of which are presented in Table 3.

Table 3

Mann-Whitney U Test for the TAM Components and TAM Total across Gender

Dimensions	**Total N**	**Mann-Whitney U**	**Wilcoxon W**	**Test Statistic**	**Standard Error**	**Standardized Test Statistic**	**Asymptotic Sig.(2-sided test)**
PEU	220	5668.5	14183.5	5668.5	461.923	-0.393	0.694
PU	220	5274.5	13789.5	5274.5	462.72	-1.244	0.214
Attitude	220	5391	13906	5391	462.358	-0.993	0.321
BI	220	5075	13590	5075	462.428	-1.676	0.094
AU	220	4795.5	13310.5	4795.5	462.425	-2.28	0.023
PE	220	5364	13879	5364	459.998	-1.057	0.291
FC	220	4996.5	13511.5	4996.5	463.483	-1.841	0.066
TC	220	5561	14076	5561	462.694	-0.625	0.532
Total TAM	220	4893	13408	4893	464.114	-2.062	0.039

As indicated in Table 3, the Mann-Whitney U tests revealed that the score distributions for male and female participants were not significantly different for the majority of TAM components, including PEU (U = 5668.5, z = -0.393, p = .694), Perceived Usefulness (U = 5274.5, z = -1.244, p = .214), attitude (U = 5391, z = -0.993, p = .321), Perceived Enjoyment (U = 5364, z = -1.057, p = .291), and Technological Complexity (U = 5561, z = -0.625, p = .532). The tests for Behavioral Intention (U = 5075, z = -1.676, p = .094) and Facilitating Conditions (U = 4996.5, z = -1.841, p = .066) approached significance, but did not exceed the alpha threshold of .05. However, a statistically significant difference was found for Actual Usage (AU) (U = 4795.5, z = -2.28, p = .023) and TAM Total score (U = 4893, z = -2.062, p = .039), with males reporting higher levels of actual use and a more positive overall perception across the model.

Addressing the Third Research Question

Before running inferential tests to compare the two groups, it was essential to inspect the descriptive statistics and test the assumption of normality for each TAM component separately for participants with and without prior experience. Table 4 depicts the descriptive statistics and normality test results for TAM components and TAM total across technology experience.

Table 4

Descriptive Statistics and Normality Test Results for TAM Components and TAM Total across Experience of Technology

TAM Components	Groups	No	Mean	SD.	Var	Kolmogorov-Smirnova			Shapiro-Wilk		
						Statistic	df	Sig.	Statistic	df	Sig.
PEU Mean	With Experience	168	3.81	.935	.87	.13	168	.00	.92	168	.00
	Without Experience	52	3.48	.96	.93	.14	52	.01	.94	52	.01
PU Mean	With Experience	168	3.83	.87	.76	.09	168	.00	.92	168	.00
	Without Experience	52	3.58	.92	.853	.15	52	.00	.95	52	.04
Attitude Mean	With Experience	168	3.75	.91	.82	.11	168	.00	.93	168	.00
	Without Experience	52	3.43	.93	.86	.11	52	.08	.96	52	.14
BI Mean	With Experience	168	3.71	.88	.78	.09	168	.00	.94	168	.00
	Without Experience	52	3.38	1.00	1.01	.10	52	.18	.96	52	.12
AU Mean	With Experience	168	3.69	.95	0.90	.12	168	.00	.93	168	.00
	Without Experience	52	3.23	1.15	1.32	.15	52	.00	.92	52	.00
PE Mean	With Experience	168	3.71	.97	0.95	.14	168	.00	.92	168	.00
	Without Experience	52	3.30	1.01	1.03	.15	52	.00	.94	52	.02
FC Mean	With Experience	168	3.51	.82	0.68	.10	168	.00	.97	168	.00
	Without Experience	52	2.94	.89	0.80	.10	52	.200*	.97	52	.42
TC Mean	With Experience	168	2.79	1.12	1.25	.11	168	.00	.94	168	.00
	Without Experience	52	2.67	1.00	1.00	.09	52	.200*	.96	52	.15
Total TAM	With Experience	168	3.61	.68	.47	.13	168	.00	.92	168	.00
	Without Experience	52	3.24	.73	.53	.09	52	.200*	.97	52	.48

As shown in Table 4, participants with prior technology experience reported higher scores on every single TAM component and total TAM compared to those without such experience. The differences were substantial for many variables, such as Facilitating Conditions (With: M = 3.51; Without: M = 2.94) and Actual Usage (With: M = 3.69; Without: M = 3.23). The results of the Kolmogorov-

Smirnov and Shapiro-Wilk tests indicate that the data for the "With Experience" group significantly deviated from normality for all components as well as the total TAM ($p < .05$). For the "Without Experience" group, the data for Attitude, BI, FC, TC, and TAM Total were normally distributed ($p > .05$), but not for others. The inspection of normality data reveals the violation of normality for all comparison pairs. Therefore, non-parametric testing was deemed the most appropriate method for analysis. Table 5 displays the Mann-Whitney U Test results for the TAM components and TAM total scores across the experience of technology groups.

Table 5

Mann-Whitney U Test for the TAM Components and TAM Total across Experience of Technology

Dimensions	Total N	Mann-Whitney U	Wilcoxon W	Test Statistic	Standard Error	Standardized Test Statistic	Asymptotic Sig.(2-sided test)
PEU	220	3485.5	4863.5	3485.5	399.14	-2.21	0.03
PU	220	3684.5	5062.5	3684.5	399.83	-1.70	0.09
Attitude	220	3455	4833	3455	399.52	-2.28	0.02
BI	220	3564	4942	3564	399.58	-2.01	0.04
AU	220	3415.5	4793.5	3415.5	399.58	-2.38	0.02
PE	220	3358.5	4736.5	3358.5	397.48	-2.54	0.01
FC	220	2735.5	4113.5	2735.5	400.49	-4.07	0.00
TC	220	3997	5375	3997	399.81	-0.92	0.35
Total TAM	220	3095.5	4473.5	3095.5	401.05	-3.17	0.00

The results of the Mann-Whitney U tests, presented in Table 5, show that there were statistically significant differences between the two groups in most TAM components. Significant differences were found for Perceived Ease of Use (U = 3485.5, z = -2.21, p = .03), attitude (U = 3455, z = -2.28, p = .02), Behavioral Intention (U = 3564, z = -2.01, p = .04), Actual Usage (U = 3415.5, z = -2.38, p = .02), Perceived Enjoyment (U = 3358.5, z = -2.54, p = .01), and Facilitating Conditions (U = 2735.5, z = -4.07, $p < .001$). Furthermore, the total TAM score was significantly higher in the experienced group (U = 3095.5, z = -3.17, p = .002). No significant difference was found for Perceived Usefulness (U = 3684.5, z = -1.70, p = .09) or Technological Complexity (U = 3997, z = -0.92, p = .35). This confirms that prior technology experience is strongly associated with a more positive acceptance profile for ChatGPT.

Qualitative Data Analysis

Addressing the Fourth Research Question

Thematic analysis of the interviews provided an explanatory context for the survey results. Table 6 presents the demographic profiles of the interview participants.

Table 6

Demographic Profile of Interview Participants

Participant ID	Gender	Year of Study	Self-Reported Proficiency	ChatGPT Usage Frequency
S1, S2, S3	Male	4th Year	Advanced	Daily
S4, S5	Female	4th Year	Advanced	Daily
S6, S7	Male	3rd Year	Upper-Intermediate	Weekly
S8, S9, S10	Female	3rd Year	Upper-Intermediate	Weekly
S11, S12	Male	2nd Year	Intermediate	Weekly
S13, S14, S15	Female	2nd Year	Intermediate	Monthly
S16, S17	Male	1st Year	Intermediate	Rare
S18, S19, S20	Female	1st Year	Pre-Intermediate	Rare

The audio-recorded interviews were transcribed verbatim and analyzed thematically following Braun and Clarke's (2006) inductive, data-driven approach. NVivo 12 software was used to assist with systematic coding and theme management. For confidentiality, pseudonyms (S1–S20) were used in the interview quotes.

Comprehensive explanations for these survey results were provided through semi-structured interviews. Twenty participants were purposively selected from a larger cohort for the interviews (N=220; 90 males, 130 females). Analysis of the interview transcripts revealed several themes related to their experiences with ChatGPT and their resulting perceptions. The themes presented below largely corroborate the survey findings, offering a balanced perspective on the advantages and disadvantages of using this tool for language learning.

Theme 1: Perceived Usefulness as a Multifaceted Learning Aid. The survey scores showed high levels of perceived usefulness (PU) for ChatGPT, which participants described as a useful and helpful pedagogical tool. Its usefulness is manifested in several ways. First, students perceived ChatGPT as a private tutor to enhance their language skills. As one learner said, "ChatGPT can check my answers and note the spelling mistakes, suggest a better way to write it, and suggest different words" (S5). Many students feel that it develops their writing and speaking skills. One student commented, "I like that it can converse with me and fix my grammatical errors, which I think will eventually help me develop my L2 skills" (S6).

Second, usefulness was strongly related to efficiency and performance. Students highlighted that the tool saved them a lot of time and effort, which was crucial to their demanding academic lifestyle. According to one participant: "The technology helps me to find the information easily which saves my time and efforts. Saving time is important for me as a college student because I would have more time to focus on other things such as analyzing, interpreting, and thinking critically." (S1) Aligning with the high PU scores, students described ChatGPT as a necessary substitute for limited teacher availability. S5 remarked, "ChatGPT can check my answers and note spelling mistakes... it suggests different words." S6 added, "While real language partners sleep and get busy, ChatGPT never does." This theme highlights how AI fills a pedagogical void in the context of Iraqi as a surrogate tutor.

Theme 2: High Perceived Ease of Use and Accessibility. The interviews were consistent with the survey results, confirming that Perceived Ease of Use (PEU) is a major determinant of students' adoption of ChatGPT. The tool's usability was rated unanimously as easy to use, intuitive, and no special training was required. One student explained, "I think I am a user of ChatGPT because it is easy to use. I do not have to develop my IT skills to use or to attend any training to be able to use it' (S7). Another participant said, "I like ChatGPT because it is easy once you logged in you can start using it" (S10). The students valued the fact that they could use it anytime, anywhere easily, as it is ubiquitous. As one learner noted, "While real language partners sleep and get busy, ChatGPT never does. "(S6).

Theme 3: Enhanced Engagement, Motivation, and Positive Attitudes. According to the responses, the ChatGPT also helped in improving engagement and motivation backed by positive attitudes. According to the students, the tool was interesting as it was interactive and personalized. Being able to role-play real-life situations was an advantage, as one student highlighted: 'When you play a job interview or a debate about a hot topic, learning grammar and building vocabulary feel like an exciting journey, which I find motivating and engaging' (S20).

They believed that by adding this knowledge to their communication and adaptation, their confidence could be enhanced and cross-cultural communication could achieve success. One student stated, "I feel convenient to talk with ChatGPT. It works quite easy on me." I do not worry about making mistakes; I love it because I feel my speaking skills are getting better." (S 17). The students were motivated to take risks and practice more freely in this non-judgmental space.

Theme 4: Drawbacks and Technological Limitations. The participants identified several major weaknesses. A major concern is accuracy and reliability. In this context, students perceived AI as lacking human contextualization. This could also create the risk of misunderstood feedback or misinterpretations. This caused them to "doubt my knowledge' (S11). The findings of the survey highlighted technological complexity as a major hurdle.

One of the limitations of this study is its text-based nature. According to the learners, ChatGPT conversations were not as interactive as human interaction. As one student put it, "ChatGPT is a great way to practice, but it doesn't quite work the same as in real life where people go back and forth." (S4) Lack of interaction can be uncomfortable. One student mentioned "It makes it hard for me to learn how to start talking, ask follow-up questions and read body language" (S18). The absence of real-time, genuine conversation was viewed as a huge disadvantage. Students were acutely aware of AI's limitations in capturing cultural and pragmatic nuances. S4 noted, "It doesn't quite work the same as in real life... it lacks body language and true interaction."

Theme 5: Concerns Regarding Over-Reliance and Academic Integrity. The interviews highlighted the concern that motivation, self-regulation, and ethics may suffer in connection with technology. Some students also believed that ChatGPT's assistance created overreliance, which contributed to their lack of ability to write and solve problems. One student admitted, "I am worried because using ChatGPT too much will undermine my writing ability, sometimes I feel unable to write without using ChatGPT (S8)."

This fear has extended to the issue of academic misconduct. The participants were conscious of the urge to copy-paste. One person said, I'm scared of using ChatGPT especially while doing my assignment because I think I might cheat (S16). Students have also raised more general ethical issues surrounding ChatGPT's generation and the spread of misinformation. It was stated by one participant that "Since Chat GPT can produce textual content that appears real and sounds real, it could produce deep fakes and spread misinformation. I have to be especially vigilant about verifying the information I receive through it' (S20).

Overall, according to the findings, Iraqi EFL learners view ChatGPT as a tool that is highly useful and easy to use, and enhances their motivation by providing them with valuable support for learning. Nonetheless, they remain aware of its limitations in terms of accuracy and conversational realism, as well as the potential risks associated with overreliance and plagiarism. Students' statements indicated that they perceived technology as an effective study friend. However, they also acknowledge that ChatGPT is imperfect and occasionally provides incorrect prompts.

Discussion

This study examined the factors governing Iraqi EFL learners' acceptance of ChatGPT using an extended Technology Acceptance Model (TAM) framework to integrate quantitative survey data with qualitative interview insights. The findings reveal a generally positive disposition towards ChatGPT, primarily driven by high Perceived Usefulness (PU) and Perceived Ease of Use (PEU). However, this acceptance is significantly mediated by learners' prior technology experience and constrained by perceived technological complexity and inadequate facilitating conditions.

Interpreting the Hierarchy of TAM Perceptions

The initial descriptive analysis showed that students prioritized the different categories of their perceptions. ChatGPT is rated as an appealing and useful tool by learners. Its usefulness (M=3.77) and ease of use (M=3.73) are the two highest-rated constructs. Studies show that technology that is both useful and easy to use creates positive attitudes and behaviors (Belda-Medina & Calvo-Ferrer, 2022; Liu et al., 2023). Afterwards, Attitude (M = 3.67), Behavioral Intention (BI) (M = 3.64), Perceived Enjoyment (PE) (M = 3.61), and Actual Usage (AU) (M = 3.58) cluster near the midpoint of the scale. This suggests the existence of a learner community that is both eager and motivated, yet it has not translated their positive perceptions into frequent utilization.

In particular, the lower end of the hierarchy is revealed. Facilitating Conditions (FC) ranked second lowest as indicated by the mean score (M=3.38), while Technological Complexity (TC) received the lowest rating with the mean score (M=2.76), which shows that this was the most important perceived barrier. This finding is significant in the Iraqi context as students are aware of the importance of ChatGPT; however, they do not use it frequently due to inconsistent institutional support, lack of training, and flaws in the tool itself. Students found the impact of ChatGPT hallucinations, issues connecting to ChatGPT, and difficulty with prompts to be the main concerns mentioned in the interviews.

The discrepancy between high Behavioral Intention and moderate Actual Use points to a specific "Technological-Pedagogical Gap" in Iraq. While students perceive the tool as useful (PU) and easy (PEU), the Facilitating Conditions, which include Internet stability, access to devices, and institutional support, lag behind. Facilitating conditions are often assumed in digitally mature contexts; in Iraq, they are a variable that actively mediates adoption. The low score for Technological Complexity indicates that the barrier is not the tool itself, but the ecosystem surrounding it

Gender Differences

The second research question explored gender differences. The analysis revealed no statistically significant differences between males and females concerning the core Technology Acceptance Model (TAM) perceptual constructs: Perceived Ease of Use (PEU), Perceived Usefulness (PU), Attitude, Perceived Enjoyment (PE), and Trust in Technology (TC), as determined by Mann-Whitney U tests. However, males reported a higher level of Actual Usage ($U = 4795.5$, $p = .023$) and a higher TAM Total Score ($U = 4893$, $p = .039$), with both differences reaching statistical significance. The tests for Behavioral Intention (BI) ($p = .094$) and Facilitating Conditions (FC) ($p = .066$) did not achieve statistical significance, although they exhibited a similar trend.

This suggests that, while both males and females think the tool is worthwhile and easy to use, they use it differently. The current study finding contrasts with studies suggesting no gender effects (e.g., Grassini et al., 2024) but is consistent with research indicating differences in gender use profiles (i.e., Belda-Medina & Calvo-Ferrer, 2022). Importantly, our interview data did not imply that this was due to gender differences in technology preferences. On the contrary, learners identified external factors such as social norms, unequal access to devices, and weak Internet access, which directly led to the second major finding of the study.

Our finding that gender does not affect attitude but does affect actual use invites a sociocultural interpretation. In the conservative and often gender-segregated context of parts of Iraq, male students may have greater access to personal digital devices and less scrutiny of their time use, thus facilitating higher actual usage. Conversely, female students in the interviews expressed higher anxiety about "rule-breaking" and plagiarism, suggesting that their lower usage may be a result of higher self-regulation and adherence to perceived institutional norms. However, some female participants also noted that ChatGPT offers a "safe space" to practice English without the social anxiety of speaking in front of male peers, indicating a complex, dual role of AI for female learners.

Prior Technology Experience

This study revealed that prior experience with technology plays a large role in differentiating services. What learners with experience reported was significantly more positive perceptions regarding all TAM components: PEU ($p = .03$, attitudee ($p = .02$, BI ($p =. 04$, AU ($p = .02$, and PE ($p = .01$). The greatest difference observed was for Facilitating Conditions ($U = 2735.5$, $z = -4.07$, $p < .001$), and as a result, the Total TAM score was also significantly higher ($p = .002$) for the experienced group.

There was no difference in PU or TC. Both groups perceived ChatGPT as equally beneficial for all users. However, the experienced group found the system more user-friendly, exhibited more favorable attitudes towards it, and demonstrated greater support, which consequently led to increased intention to use and actual usage. To effectively transform value into practical application, it is essential to possess prior experience with technology and readiness to engage with it. Qualitative data indicate that users with experience in utilizing this tool employ advanced prompting strategies and demonstrate a superior ability to navigate its complexities.

The strong influence of prior experience on all constructs, especially Facilitating Conditions, suggests that digital literacy acts as a form of capital. Experienced users possess the "troubleshooting resilience" to navigate Iraq's spotty Internet and access AI tools effectively. They perceived the same environment as more facilitating than their novice peers ($p <.001$), proving that "access" is not just about infrastructure but also about the user's capability to exploit that infrastructure.

Learners' Perspectives

The results of this qualitative analysis indicate that Iraqi EFL learners view ChatGPT as a versatile personal tutor. ChatGPT is presumed to positively influence writing and speaking, while saving time and sparking interest. At the same time, the respondents believe that ChatGPT has some risks, including the accuracy of information provided, ability to have dialogues, over-reliance, and academic honesty.

Technology acceptance theories closely align with the observed results. Specifically, perceived usefulness and ease of use are fundamental to acceptance (Davis, 1989). Moreover, these factors have been demonstrated to influence the adoption of AI chatbots by EFL and EFL/ESL learners (Qu and Wu 2024). The emphasis students place on ease of use and 24/7 accessibility mirrors previous research on chatbot partners for language learning, as well as the concept of ubiquitous, low-friction practice (Fryer et al., 2019).

Learners have reported reduced communication anxiety and greater confidence, which echoes the early L2 literature. The findings highlight low-stakes practice affordances (Fryer et al., 2019) and recent accounts of ChatGPT's capacity to scaffold language production and writing processes (Kohnke et al., 2023; Song & Song, 2023). Participants' concerns about accuracy, reliability, and the need for critical evaluation reflect the documented risks of hallucination. They also refer to misinformation and context insensitivity, which are more relevant to ChatGPT-type systems (Kasneci et al., 2023). As demonstrated by studies calling attention to the lack of empirical pragmatic sophistication and cultural sensitivity of AI in EFL writing and interaction (Hwang et al., 2025), there is a perceived gap between chatbot exchanges and genuine reciprocal human conversation. In conclusion, students' ambivalence regarding over-reliance and academic integrity reflects emerging evidence in higher education: learners seek efficiency but express concerns about the erosion of authorship, temptation to plagiarize, and ambiguous boundaries of responsible use (Johnston et al., 2024; Kasneci et al., 2023). The qualitative data revealed that students are operating in a "policy vacuum." The Ministry of Higher Education has yet to issue comprehensive guidelines on Generative AI. Consequently, students are left to construct their own ethical frameworks, oscillating between seeing ChatGPT as a helpful "tutor" and a tool for "cheating." This ambiguity created the anxiety observed in Theme 5, highlighting the urgent need for institutional guidance.

Conclusion

The findings of this study show that Iraqi EFL learners view ChatGPT as a beneficial, "surrogate tutor" that enhances proficiency and motivation. However, effective implementation is stratified by gender and digital experience and is hampered by infrastructural limitations. These findings have several pedagogical implications. It is recommended that universities implement foundational digital literacy workshops ("AI Clinics") to scaffold digital literacy. They should also initiate structured onboarding processes for students with limited experience to ensure equitable access to technology. These should focus on prompt engineering and critical evaluation of AI outputs to bridge the gap between experienced and novice users. The construct of FC was identified as having a low score, suggesting that it serves as an optimal institutional investment differentiator among experience groups. The research findings propose solutions such as ensuring Internet availability, formulating an ethical use policy, and providing training on effective prompting, which can directly address these barriers. While students acknowledged the utility of ChatGPT, they also recognized the risks associated with inaccuracies and over-reliance. Consequently, educational curricula must evolve to include instruction on verifying AI-generated messages, validating information, and utilizing the tool as an aid, rather than as a substitute for developing core language skills. The Ministry of Higher Education must formulate clear bilingual (Arabic/English) policies that define responsible AI use. This will alleviate student anxiety regarding plagiarism and legitimize the tool as a learning aid. Administrators must prioritize reliable Internet access in campus libraries to ensure equity for students who lack "facilitating conditions" at home. The study relied on convenience sampling, which may have overrepresented students with better digital access. Future research should employ longitudinal designs to track adoption over time and use Structural Equation Modeling (SEM) with larger, representative samples to map causal pathways more precisely.

References

Abdaljaleel, M., Barakat, M., Alsanafi, M., Salim, N. A., Abazid, H., Malaeb, D., ... & Sallam, M. (2024). A multinational study on the factors influencing university students' attitudes and usage of ChatGPT. *Scientific Reports*, *14*(1), 1983. https://doi.org/10.1038/s41598-024-52549-8

Alharbi, S., & Drew, S. (2014). Using the technology acceptance model in understanding academics' behavioural intention to use learning management systems. *International Journal of Advanced Computer Science and Applications, 5*(1), 143-155. https://dx.doi.org/10.14569/IJACSA.2014.050120

Ali, J. K. M., Shamsan, M. A. A., Hezam, T. A., & Mohammed, A. A. (2023). Impact of ChatGPT on learning motivation: Teachers and students' voices. *Journal of English Studies in Arabia Felix*, *2*(1), 41–49. https://doi.org/10.56540/jesaf.v2i1.51

Al-Nuaimi, M. N., & Al-Emran, M. (2021). Learning management systems and technology acceptance models: A systematic review. *Education and Information Technologies*, *26*(5), 5499-5533. https://doi.org/10.1007/s10639-021-10513-3

Alotaibi, H. M., Sonbul, S. S., & El-Dakhs, D. A. (2025). Factors influencing the acceptance and use of ChatGPT among English as a foreign language learners in Saudi Arabia. *Humanities and Social Sciences Communications*, *12*, 628. https://doi.org/10.1057/s41599-025-04945-2

Alotumi, M. (2022). Factors influencing graduate students' behavioral intention to use Google Classroom: Case study-mixed methods research. *Education and Information Technologies*, *27*(7), 10035-10063. https://doi.org/10.1007/s10639-022-11051-2

Alshahrani, A., Dennehy, D., & Mäntymäki, M. (2022). An attention-based view of AI assimilation in public sector organizations: The case of Saudi Arabia. *Government Information Quarterly*, *39*(4), 101617. https://doi.org/10.1016/j.giq.2021.101617

Ameen, N., Tarhini, A., Reppel, A., & Anand, A. (2021). Customer experiences in the age of artificial intelligence. *Computers in Human Behavior*, *114*, 106548. https://doi.org/10.1016/j.chb.2020.106548

Bai, B., Wang, J., & Zhou, H. (2022). An intervention study to improve primary school students' self-regulated strategy use in English writing through e-learning in Hong Kong. *Computer Assisted Language Learning*, *35*(9), 2265-2290. https://doi.org/10.1080/09588221.2020.1871030

Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using chatbots as AI conversational partners in language learning. *Applied Sciences*, *12*(17), 8427. https://doi.org/10.3390/app12178427

Bin-Hady, W. R. A., Al-Kadi, A., Hazaea, A., & Ali, J. K. M. (2025). Exploring the dimensions of ChatGPT in English language learning: A global perspective. *Library Hi Tech*. https://doi.org/10.1108/LHT-05-2023-0200

Bodani, N., Lal, A., Maqsood, A., Altamash, S., Ahmed, N., & Heboyan, A. (2023). Knowledge, attitude, and practices of general population toward utilizing ChatGPT: A cross-sectional study. *SAGE Open*, *13*(4), 21582440231211079. https://doi.org/10.1177/21582440231211079

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Chuttur M.Y. (2009). "Overview of the Technology Acceptance Model: Origins, Developments and Future Directions," Indiana University, USA. *Sprouts: Working Papers on Information Systems*, 9(37). https://aisel.aisnet.org/sprouts_all/290

Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, *61*(2), 228-239. https://doi.org/10.1080/14703297.2023.2190148

Creswell, J. W., & Clark, V. P. (2007). *Mixed methods research*. Thousand Oaks, CA.

Davis, F. D. (1986). *A technology acceptance model for empirically testing new end-user information systems: Theory and results* (Doctoral dissertation, Massachusetts Institute of Technology).

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3). 319-340. https://doi.org/10.2307/249008

Dennehy, D., Schmarzo, B., & Sidaoui, M. (2022). Organising for AI-powered innovation through design: the case of Hitachi Vantara. *International Journal of Technology Management*, *88*(2-4), 312-334. https://doi.org/10.1504/IJTM.2022.121507

Draxler, F., Buschek, D., Tavast, M., Hämäläinen, P., Schmidt, A., Kulshrestha, J., & Welsch, R. (2023). Gender, age, and technology education influence the adoption and appropriation of LLMs. arXiv preprint arXiv: 2310.06556. https://doi.org/10.48550/arXiv.2310.06556

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., ... & Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *57*, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Elshaer, I. A., Hasanein, A. M., & Sobaih, A. E. E. (2024). The moderating effects of gender and study discipline in the relationship between university students' acceptance and use of ChatGPT. *European Journal of Investigation in Health, Psychology and Education*, *14*(7), 1981–1995. https://doi.org/10.3390/ejihpe14070132

Farr, C. (2024). *Unmasking ChatGPT: The Challenges of Using Artificial Intelligence for Learning Vocabulary in English as an Additional Language* (Doctoral Dissertation).

Fryer, L. K., Nakao, K., & Thompson, A. (2019). Chatbot learning partners: Connecting learning experiences, interest and competence. *Computers in Human Behavior*, *93*, 279-289. https://doi.org/10.1016/j.chb.2018.12.023

Granić, A. and Marangunić, N. (2019). Technology acceptance model in educational context: a systematic literature review. *British Journal of Educational Technology*, *50*(5), 2572-2593. https://doi.org/10.1111/bjet.12864

Grassini, S., Aasen, M. L., & Møgelvang, A. (2024). Understanding University Students' Acceptance of ChatGPT: Insights from the UTAUT2 Model. *Applied Artificial Intelligence*, *38*(1), 2371168. https://doi.org/10.1080/08839514.2024.2371168

Han, J., Yoo, H., Kim, Y., Myung, J., Kim, M., Lim, H., ... & Oh, A. (2023). RECIPE: How to Integrate ChatGPT into EFL Writing Education. In Proceedings of the Tenth ACM Conference on Learning @ Scale, (416-420). https://doi.org/10.48550/arXiv.2305.11583

Hong, W. C. H. (2023). The impact of ChatGPT on foreign language teaching and learning: opportunities in education and research. *Journal of Educational Technology and Innovation*, *5*(1). 35-45. https://doi.org/10.61414/jeti.v5i1.103

Hwang, M., Lee, E., & Lee, H. K. (2025). Exploring EFL learners' acceptance of ChatGPT: Application of the extended technology acceptance model. *English Teaching*, *80*(1), 45–69. https://doi.org/10.15858/engtea.80.1.202503.45

Javaid, M., Haleem, A., Singh, R. P., Khan, S., & Khan, I. H. (2023). Unlocking the opportunities through ChatGPT Tool towards ameliorating the education system. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, *3*(2), 1-12. https://doi.org/10.1016/j.tbench.2023.100115

Jeon, J., Lee, S., & Choi, S. (2023). A systematic review of research on speech-recognition chatbots for language learning: Implications for future directions in the era of large language models. *Interactive Learning Environments*, 1-19. https://doi.org/10.1080/10494820.2023.2204343

Johnston, H., Wells, R. F., Shanks, E. M., Boey, T., & Parsons, B. N. (2024). Student perspectives on the use of generative artificial intelligence technologies in higher education. *International Journal for Educational Integrity*, *20*(1), 2. https://doi.org/10.1007/s40979-024-00149-4

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Khanchel, H. (2023). Factors affecting social network use by students in Tunisia. *Human Systems Management*, *42*(2), 131-148. https://doi.org/10.3233/hsm-220017

Kim, S., Shim, J., & Shim, J. (2023). A Study on the Utilization of OpenAI ChatGPT as a Second Language Learning Tool. *Journal of Multimedia Information System*, *10*(1), 79-88. https://doi.org/10.33851/JMIS.2023.10.1.79

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54 (2). 1-14. https://doi.org/10.1177/00336882231162868

Koraishi, O. (2023). Teaching English in the age of AI: Embracing ChatGPT to optimize EFL materials and assessment. *Language Education and Technology*, *3*(1). 55-72. https://langedutech.com/letjournal/index.php/let/article/view/48/37

Kostka, I., & Toncelli, R. (2023). Exploring applications of ChatGPT to English language teaching: Opportunities, challenges, and recommendations. *TESL-EJ*, *27*(3). 1-19. https://www.tesl-ej.org/pdf/ej107/int.pdf

Kotsis, K. T. (2024). Artificial Intelligence Creates Plagiarism or Academic Research?. *European Journal of Arts, Humanities and Social Sciences*, *1*(6), 169-179. https://doi.org/10.59324/ejahss.2024.1(6).18

Li, X., Li, B., & Cho, S. J. (2023). Empowering Chinese Language Learners from Low-Income Families to Improve Their Chinese Writing with ChatGPT's Assistance Afterschool. *Languages*, *8*(4), 238. https://doi.org/10.3390/languages8040238

Liu, G. L., Darvin, R., & Ma, C. (2024). Exploring AI-mediated informal digital learning of English (AI-IDLE): a mixed-method investigation of Chinese EFL learners' AI adoption and experiences. *Computer Assisted Language Learning*, 1-29. https://doi.org/10.1080/09588221.2024.2310288

Ma, J., Wang, P., Li, B., Wang, T., Pang, X. S., & Wang, D. (2024). Exploring User Adoption of ChatGPT: A Technology Acceptance Model Perspective. *International Journal of Human–Computer Interaction*, 1-15. https://doi.org/10.1080/10447318.2024.2314358

Mhlanga, D. (2023). Open AI in Education, the Responsible and Ethical Use of ChatGPT Towards Lifelong Learning. *Social Science Research Network, 2023,* 4354422. https://doi.org/10.2139/ssrn.4354422

Nadri, H., Rahimi, B., Afshar, H. L., Samadbeik, M., & Garavand, A. (2018). Factors affecting acceptance of hospital information systems based on extended technology acceptance model: a case study in three paraclinical departments. *Applied Clinical Informatics*, *9*(02), 238-247. https://doi.org/10.1055/s-0038-1641595

Panagiotarou, A., Stamatiou, Y. C., Pierrakeas, C., & Kameas, A. (2020). Gamification acceptance for learners with different e-skills. *International Journal of Learning, Teaching and Educational Research*, *19*(2), 263–278. https://doi.org/10.26803/ijlter.19.2.16

Porto, N., Huang, Y., & Xiao, J. J. (2019). Credit card adoption and usage in China: urban–rural comparisons. *The Singapore Economic Review*, *64*(01), 41-56. https://doi.org/10.1142/S021759081743010X

Punar Özçelik, N., & Yangın Ekşi, G. (2024). Cultivating writing skills: the role of ChatGPT as a learning assistant—a case study. *Smart Learning Environments*, *11*(1), 10. https://doi.org/10.1186/s40561-024-00296-8

Qu, K., & Wu, X. (2024). ChatGPT as a CALL tool in language education: A study of hedonic motivation adoption models in English learning environments. *Education and Information Technologies*, *29*(15), 19471-19503. https://link.springer.com/article/10.1007/s10639-024-12598-y

Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences*, *13*(9), 5783. https://doi.org/10.3390/app13095783

Sallam, M. (2023, March). ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare 11*(6), 1-34. https://doi.org/10.3390/healthcare11060887

Schoonenboom, J. (2012). The use of technology as one of the possible means of performing instructor tasks: Putting technology acceptance in context. *Computers & Education, 59*(4), 1309–1316. https://doi.org/10.1016/j.compedu.2012.06.009

Shiri, A. (2023). ChatGPT and academic integrity. *Information Matters*, *3*(2). 1-11. https://informationmatters.org/2023/02/chatgpt-and-academic-integrity/

Sindermann, C., Yang, H., Elhai, J., Yang, S., Ling, Q., Mei, L., … & Montag, C. (2022). Acceptance and fear of artificial intelligence: associations with personality in a German and a Chinese sample. *Discover Psychology*, *2*(1): 8. https://doi.org/10.1007/s44202-022-00020-y

Solano-Lorente, M., Martínez-Caro, E., & Cegarra-Navarro, J. G. (2013). Designing a framework to develop eLoyalty for online healthcare services. *Electronic Journal of Knowledge Management*, *11*(1), 107-115.

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, *14*, 1260843. https://doi.org/10.3389/fpsyg.2023.1260843

Tandon, A., Sharma, H., & Aggarwal, A. G. (2019). Studying the moderating effect of a respondent's locality in M-commerce adoption intention. *Ingeniería Solidaria*, *15*(3), 1-23. https://doi.org/10.16925/2357-6014.2019.03.07

Tarhini, A., Hone, K., & Liu, X. (2015). A cross-cultural examination of the impact of social, organisational and individual factors on educational technology acceptance between British and Lebanese university students. *British Journal of Educational Technology*, *46*(4), 739-755. https://doi.org/10.1111/bjet.12169

Utebayeva, G. (2024). *Learning English in the era of artificial intelligence: Graduate students' experiences and perceptions of ChatGPT at an EMI university in Kazakhstan. Suleyman Demirel University* (Unpublished Dissertation).

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478. https://doi.org/10.2307/30036540

Wang, T., Lund, B., Marengo, A., Pagano, A., Mannuru, N., Teel, Z., … & Pange, J. (2023). Exploring the potential impact of artificial intelligence (ai) on international students in higher education: generative ai, chatbots, analytics, and international student success. *Applied Sciences*, *13*(11), 6716. https://doi.org/10.3390/app13116716

Weidener, L. and Fischer, M. (2023). Teaching AI ethics in medical education: a scoping review of current literature and practices. *Perspectives on Medical Education*, 12(1), 399-410. https://doi.org/10.5334/pme.954

Xiao, Y., & Zhi, Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. *Languages*, *8*(3), 212. https://doi.org/10.3390/languages8030212

Xiong, H. (2024). Research on confusing responses based on ChatGPT. *Applied and Computational Engineering*, 57, 90-97. https://doi.org/10.54254/2755-2721/57/20241315

Zhang, B., Zhu, Y., Deng, J., Zheng, W., Liu, Y., Wang, C., & Zeng, R. (2023). "I am here to assist your tourism": Predicting continuance intention to use AI-based chatbots for tourism. Does gender really matter? *International Journal of Human–Computer Interaction*, *39*(9), 1887–1903. https://doi.org/10.1080/10447318.2022.2124345

Zhuo, T. Y., Huang, Y., Chen, C., & Xing, Z. (2023). Exploring ai ethics of ChatGPT: A diagnostic analysis. *arXiv preprint arXiv:2301.12867*. https://doi.org/10.48550/arXiv.2301.12867

Appendix A: Online Survey

Here is the link to the online survey

https://docs.google.com/forms/d/e/1FAIpQLSfDA7Stu8x5bpKhLvLamPlxBtpUr019-A1zyEoMuySJBPS6lA/viewform

Appendix B: Interview Guiding Questions

1. How familiar are you with ChatGPT? Have you ever used it for language learning?
2. What are your overall impressions of using ChatGPT for language learning?
3. What are the biggest benefits of using ChatGPT for language learning?
4. How would you compare ChatGPT and traditional language learning methods?
5. Can you describe a specific instance where ChatGPT helped you learn something new or overcome a language learning challenge?
6. Have you found any limitations in ChatGPT's ability to provide accurate or helpful information?
7. Do you ever feel frustrated or confused by ChatGPT's responses?

Authors' Biography

Ahmad Rawdan Salman is postdoctoral researcher of English language and applied linguistics at Razi University in Iran. His academic interests include educational technology, digital literacy development, and innovations in English language teaching. He has published research articles in national and international journals.

Saman Ebadi is a full professor of applied linguistics at Razi University, Kermanshah, Iran. His areas of interest are dynamic assessment, CALL, an artificial intelligence in language teaching. He has authored books and published papers in national/international conferences and journals.

Sajad Velayati is a researcher and PhD student in Applied Linguistics at Razi University, Kermanshah, Iran. His research interests include dynamic assessment, CALL, sociocultural theory, flipped instruction and the integration of artificial intelligence in language education. He has authored books and published papers in national/international conferences and journals.

Volume 14, Issue 1
Winter and Spring, 2025
pp. 123-146

Journal of Applied Linguistics and Applied Literature: Dynamics and Advances

Evaluating Listening Tasks in IELTS Coursebooks

Sasan Baleghizadeh[1*], Mehran Mousavi[2]

[1] *Corresponding author: Professor of Applied Linguistics, Department of English Language and Literature, Faculty of Letters and Human Sciences, Shahid Beheshti University, Tehran, Iran, ORCID: 0000-0002-2290-8322*
Email: sasanbaleghizadeh@yahoo.com

[2] *PhD Candidate in TEFL, Shahid Beheshti University, Tehran, Iran*
Email: mehran.mousaviii@gmail.com

Abstract

Although IELTS preparation materials often claim to promote communicative competence, there is growing concern that they prioritize test-taking strategies over the development of authentic listening skills. This study examines the communicative orientation of listening tasks in six widely used IELTS preparation coursebooks in Iran (selected based on consultations with 15 institute directors across four major cities). In this context, learners heavily rely on instructional materials due to limited access to authentic English input. Utilizing an integrated framework that synthesizes Goh's (2001) task typology with Bachman and Palmer's (1996) model of communicative language ability, the study analyzed 565 listening tasks, a large sample, drawn from these internationally published IELTS coursebooks. Tasks were classified into four categories, one-way transactional, one-way interactional, two-way transactional, and two-way interactional, based on their participatory structure and communicative purpose. The findings show a clear imbalance; tasks are predominantly one-way transactional, emphasizing passive information extraction with limited opportunities for interactive engagement or pragmatic interpretation. While some coursebooks exhibit a gradual progression toward communicative authenticity across proficiency levels, others remain heavily exam-oriented. This variation reflects inconsistent support for developing real-world listening skills. In the Iranian EFL context, such a focus on non-interactive tasks may hinder learners' pragmatic and strategic competence, leaving them underprepared for academic and professional communication. The study highlights the need for more communicatively oriented materials and pedagogical strategies aligned with national curriculum goals.

Keywords: IELTS preparation materials, listening task authenticity, communicative competence, washback effects

ARTICLE INFO
Review Article
Received: Sunday, September 28, 2025
Accepted: Sunday, March 16, 2026
Published: Wedensday, April 1, 2026
Available Online: Sunday, March, 16, 2026
DOI: https://doi.org/ 10.22049/jalda.2026.31055.1890
Online ISSN: 2821-0204; Print ISSN: 28208986

Introduction

High-stakes English proficiency tests, such as the International English Language Testing System (IELTS), have a significant impact on English as a Foreign Language (EFL) teaching practices, often leading to a greater focus on test preparation than on the development of communicative competence (Tsagari & Cheng, 2017). As a result, learners may become proficient in test-taking strategies while lacking the communicative skills necessary for academic and professional success (Cheng et al., 2004; Green, 2007; Read, 2022). This disconnect is particularly problematic in listening instruction, where authentic comprehension involves processing linguistic input, interpreting pragmatic meaning, and responding in real time, skills that require meaningful, interactive practice (Goh & Vandergrift, 2012; Rost & Brown, 2022). Nevertheless, many IELTS materials continue to prioritize discrete listening skills over the development of integrated, communicative abilities (Field, 2012; Hughes, 1989).

In Iran, this challenge is magnified by limited access to authentic English outside the classroom and a heavy reliance on commercially published IELTS preparation materials (Babaee, 2017; Saif, 2021). Despite recent curriculum reforms promoting communicative language teaching (Tajeddin & Chamani, 2020), many commonly used materials remain test-driven, offering tasks that do not align with the interactional and functional demands of real-world listening (Moharami & Daneshfar, 2022; Rasti, 2018). Since these textbooks often act as the de facto curriculum in many Iranian classrooms, they play a major role in shaping washback. Therefore, analyzing their design is not merely an evaluation of content, but a crucial step in understanding the actual mechanism of washback in everyday classroom learning.

Despite a well-established body of research on test washback (e.g., Cheng et al., 2004; Knoch et al., 2020; Saif, 2021), empirical investigations continue to rely heavily on questionnaires and interviews. As Cheng et al. (2015) note, much of the existing evidence is perception-based rather than grounded in objective analyses of instructional artifacts. A related, yet often disconnected, problem exists in research on authenticity in language assessment. A recent systematic review (Hasrol et al., 2022) indicates that while researchers often distinguish between situational and interactional authenticity, few examine the real-world Target Language Use (TLU) domains that tasks are intended to represent. The primary gap, therefore, is twofold: (1) existing washback studies rely on subjective reports rather than material analysis, and (2) the material analyses that do exist often focus on internal test-task features without systematically comparing them to real-world communicative demands.

To address these gaps, the present study adopts an integrated analytical framework that synthesizes Goh's (2001) typology of listening tasks with Bachman and Palmer's (1996) model of communicative language ability.

While neither framework is sufficient alone to capture the complexity of washback in materials, together they offer a solid, two-sided approach. Goh's typology provides a necessary structural classification, categorizing tasks by participatory structure (one-way vs. two-way) and communicative function (transactional vs. interactional). However, classification alone does not equal evaluation. Therefore, Bachman and Palmer's model is employed to assess the quality of these tasks, specifically measuring their authenticity against real-world linguistic, functional, and sociolinguistic demands. This synthesis allows for a comprehensive critique: Goh identifies the pedagogical goal, while Bachman and Palmer highlight the gap between that goal and real-world use.

The study examines six internationally published IELTS preparation coursebooks, identified through institutional surveys (n = 15) as the most widely used in Iran's private language institute sector. Through this integrated, theory-informed analysis, the study examines how current IELTS listening materials influence learners' communicative development. Ultimately, it aims to determine, based on empirical evidence, whether the communicative label used for these materials truly reflects meaningful interactional competence or a sophisticated form of test-wiseness.

Literature Review

From Theory to Practice: Shaping Authenticity in Listening Task Design

The shift from structuralist to communicative paradigms has significantly reshaped the design of listening tasks. In Communicative Language Teaching, authenticity extends beyond real-life topics to encompass communicative intent and interactional features (Field, 2023; Gilmore, 2007). It involves situational and interactional realism, unscripted language, cognitive engagement, and cultural relevance (Hasrol et al., 2022). Culturally-laden tasks offer an opportuninty for raising cultural awareness and promoting learners' intercultural communicative competence (ICC) (Badpa & Beikian, 2025; Esmaeili et al., 2022). Listening is now viewed not as passive reception, but as an active process in which learners construct meaning within a specific context (Goh & Vandergrift, 2012).

Authentic tasks support both language development and motivation. Bouftira et al. (2022) found that interactionally authentic tasks enhance pragmatic competence, particularly in blended settings, while Gilabert et al. (2009) demonstrated that tasks requiring negotiation of meaning better prepare learners for real-life communication than traditional drills (Ballesteros Soria & van Compernolle, 2023). Consequently, scholars such as Ockey and Wagner (2018) advocate for tasks utilizing real spoken texts, varied accents, and multimedia interaction, arguing that high-stakes tests like IELTS must reflect these authentic academic and professional exchanges (Hasrol et al., 2022).

However, implementing authentic materials in high-stakes contexts poses challenges. Learners may feel overwhelmed by complex texts (Peacock, 1997), and without adequate scaffolding, beginners often struggle (Nunan, 1999). Furthermore, digital formats require nuanced classification; as Goh and Vandergrift (2012) note, a video lecture may remain a one-way transactional task, whereas an online simulation may involve two-way interaction. This complexity necessitates a rigorous framework to evaluate whether materials achieve the balance between authenticity and accessibility.

Evaluating Listening Task Authenticity: Conceptual Models and Limitations

To evaluate authenticity, researchers have proposed various frameworks, though their applicability to communicative task analysis varies. Some models, while insightful, are less suited for evaluating communicative authenticity. Buck's (2001) system emphasizes input and response formats over communicative intent.

Field's (2003) cognitive framework focuses on mental processes but overlooks the social dimensions of interaction. Similarly, Robinson's (2001) task complexity framework, though practical for gauging difficulty, often blurs the line between complexity and authenticity (Brunfaut & Révész, 2015), risking the prioritization of linguistic demands over communicative relevance.

To overcome these limitations, the present study proposes an integrated approach utilizing two complementary frameworks.

At the macro level, Bachman and Palmer's (1996) framework provides the necessary criteria for establishing situational authenticity. They conceptualize authenticity as the degree of correspondence between the characteristics of a language test task and those of a Target Language Use (TLU) domain. Their model focuses on defining the external parameters of the communicative event, specifically the physical setting, the participants' roles, and the purpose of the interaction. However, while Bachman and Palmer articulate the broad conditions necessary for authentic assessment, their framework offers limited guidance for analyzing the fine-grained procedural characteristics and psycholinguistic demands of specific pedagogical activities used in the classroom.

To address these limitations, Goh's (2001) typology is employed at the micro level to categorize the internal mechanics of the listening tasks. Drawing on the functional distinctions proposed by Brown and Yule (1983), Goh classifies activities based on participatory structure (one-way vs. two-way) and communicative function (transactional vs. interactional). This classification is essential because it serves as a proxy for cognitive difficulty; for instance, 'one-way transactional' tasks place distinct demands on learner memory and processing compared to 'two-way interactional' tasks (Goh, 2000). By integrating Goh's focus on cognitive processing with Bachman and Palmer's situational constraints, this study establishes a comprehensive framework that accounts for both the social reality of the event and the psycholinguistic reality of the learner. This dual perspective is essential for ensuring construct validity, as it prevents the validation of superficially realistic tasks, activities that look like real-world tasks on the surface but fail to elicit authentic cognitive engagement.

Gaps in the Authenticity of Listening Tasks in IELTS Commercial Coursebooks

While the construct validity of the IELTS test has been widely examined (Alavi et al., 2018; Aryadoust, 2011; Phakiti, 2016), less attention has been paid to the listening tasks in IELTS preparation coursebooks. Most research has concentrated on reading and writing (Chalmers & Walkinshaw, 2014; Marina, 2018; Priyanti, 2017), leaving listening materials underexplored.

This gap is particularly consequential in EFL contexts, such as Iran, where learners rely heavily on coursebooks as their primary source of authentic input (Kirkpatrick, 2020). Research suggests that commercial IELTS materials often prioritize test-taking strategies, such as keyword matching, over situational and interactional authenticity (Huang & Zeng, 2025; Minh, 2025; Nguyen & Le, 2024). This emphasis fosters negative washback, where instruction focuses on surface-level lexical cues rather than discourse comprehension (Bailey, 1996; Green, 2007; Mickan & Motteram, 2009). Empirical studies confirm that such materials often lead to exam-driven methods that fail to foster authentic listening competence (Estaji & Ghiasvand, 2021; Wen et al., 2022).

Crucially, no previous study has combined these two specific frameworks to analyze IELTS materials. Existing research often looks at textbooks too broadly or lacks a detailed method to spot specific flaws. By using Goh's typology to classify what the tasks are, and Bachman and Palmer's model to evaluate how authentic they are, this study can identify exactly where materials fail, whether the situation is unrealistic (e.g., a fake purpose) or the mental processing is too simple (e.g., a difficult text that only requires a simple answer). To address this clear methodological and theoretical gap, this study poses the following research questions:

RQ1: What is the frequency distribution of listening tasks across Goh's (2001) four typological categories within the selected IELTS preparation materials?

RQ2: Are there statistically significant differences in the proportion of communicatively oriented tasks (two-way and interactional) across different coursebooks and proficiency levels?

RQ3: How do the observed task patterns align with stated pedagogical objectives, and what are the implications for communicative competence development?

Method

Research Design

This mixed-methods study examined listening tasks in IELTS preparation coursebooks using Goh's (2001) framework and Bachman and Palmer's model to assess alignment with communicative language teaching principles. Quantitative analysis addressed RQ1 and RQ2 by categorizing task types and frequencies, with a focus on two-way interactional tasks that develop real-world listening skills. Qualitative analysis for RQ3 assessed how task patterns align with the coursebooks' stated goals and their implications for the development of communicative competence.

Sampling and Corpus Selection

To address the requirement for representativeness in the Iranian EFL context, a two-stage sampling procedure was established.

Stage 1: Institutional Adoption Survey

A preliminary market survey was conducted across 15 major language institutes in four metropolitan centers (Tehran, Isfahan, Kerman, and Mashhad). These institutes represent the top tier of the private sector, collectively serving approximately 12,000 learners annually. Directors of Studies were surveyed regarding their current core IELTS curriculum. Inclusion criteria for the corpus were defined as:

1. Market Penetration: The title must be the primary coursebook in the surveyed institutes.

2. Publisher Status: Must be produced by a major international ELT publisher (e.g., Cambridge, Macmillan, Collins) to ensure global relevance.
3. Currency**:** Published or revised within the last 13 years (post-2012) to align with current IELTS test specifications.
4. Proficiency range: The books cover IELTS bands 4.0–7.5, from Pre-intermediate to Advanced levels.
5. Listening focus: Each coursebook must contain substantial, integral listening components rather than supplementary ones.

Stage 2: Final Corpus Selection

Based on the survey data, six coursebooks were selected. These texts span the proficiency bands 4.0–7.5, ensuring the analysis captures the progression from foundational to advanced instruction.

Table 1

Selected IELTS Preparation Coursebooks

Title	Authors	Publisher	Details
Mindset for IELTS (Foundation to Level 3)	Foundation: (Archer et al., 2017) Level 1: (Crosthwaite et al., 2017) Level 2: (De Souza, 2017) Level 3: (Archer et al., 2018)	Cambridge University Press	2017–2018; digital pack update (2023, not currently accessible in Iran)
Complete IELTS (Bands 4–5 to 6.5–7.5)	Bands 4–5: (Wyatt, 2012) Bands 5–6.5: (Brook-Hart & Jakeman, 2012) Bands 6.5–7.5: (Brook-Hart & Jakeman, 2013)	Cambridge University Press	2012–2013; new edition (July 2025, not yet available in Iran)
Collins Listening for IELTS	(Aish & Tomlinson, 2019)	HarperCollins	Revised edition, 2019
Ready for IELTS (2nd Edition)	(McCarter, 2017)	Macmillan Education	2017 (Second Edition)
IELTS Advantage: Speaking and Listening Skills	(Marks, 2017)	DELTA Publishing	Reprinted 2017
Barron's IELTS	(Lougheed, 2016)	Barron's Educational Series	2016

A total of N = 565 discrete listening tasks were extracted. A task was operationally defined as any pedagogical unit requiring the processing of spoken input (audio or video) to produce a measurable response. Mechanical drills (e.g., listen and repeat) and post-listening discussions without audio re-play were excluded.

Coding Manual and Operational Rules

To ensure reproducibility, we developed a clear coding protocol for classifying tasks using Goh's (2001) typology. The protocol followed a simple yes/no decision tree designed to reduce subjective judgments.

Operational Decision Rules

Determining Participatory Structure (One-Way vs. Two-Way):

Rule: Does the learner have any power to affect what the speaker says or how the interaction unfolds?

Code One-Way if: The input is pre-recorded or scripted, and the learner is a passive recipient (e.g., listening to a lecture, filling gaps).

Code Two-Way if: The task asks the learner to respond in real time to a stimulus that shifts based on their input, or to take part in a role-play where meaning is negotiated. (e.g., Ask your partner for clarification on the lecture points).

Determining Communicative Function (Transactional vs. Interactional)

Rule: Is the primary goal information transfer or social/pragmatic maintenance?

Code Transactional if: The focus is on message transmission, facts, or details (e.g., What time does the train leave?).

Code Interactional if: The focus is on interpersonal relationships, expressing feelings, or interpreting stance/pragmatics (e.g., How does the speaker feel about the proposal? or persuade your partner).

Analytical Framework

Theoretical framework: Goh's Typology

Goh's (2001) listening task typology served as the main framework, adapted for contemporary multimodal formats. This study made principled adaptations, which involved two phases:

- Phase 1: Expert review. Four applied linguistics experts reviewed the proposed adaptations. The panel included experienced IELTS instructors (minimum seven years teaching experience) with advanced TESOL training (a PhD candidate and three M.A. graduates in TEFL).
- Phase 2: Pilot testing. The refined framework was tested on 50 listening tasks from the initial coursebook survey. Two independent raters applied the typology, and discrepancies were resolved through discussion. This process ensured clarity and consistency before full-scale application.

Data Analysis

Quantitative Analysis

In the quantitative analysis, descriptive statistics were used to show how listening task types were distributed across the four categories: one-way transactional, one-way interactional, two-way transactional, and two-way interactional. This study calculated frequencies and percentages for each category, both overall and within each coursebook, and compared the results to see whether certain books or levels leaned more toward specific task types.

All statistical analyses were performed using IBM SPSS Statistics (Version 28). Chi-square tests of independence assessed task-type distribution differences both within series (across proficiency levels) and between series (e.g., Mindset vs. Complete IELTS). Cramér's V complemented significant χ^2 results by quantifying effect sizes, interpreted as: negligible (<0.10), weak (0.10–0.29), moderate (0.30–0.49), or strong (≥0.50).

Alignment with Stated Objectives

To address RQ3, this study evaluated how listening task distributions aligned with each coursebook's stated pedagogical objectives using Bachman and Palmer's (1996) communicative competence model. The analysis involved three stages:

Systematic content analysis: All coursebook introductions and teacher guides were systematically coded using *Atlas.ti 9.0* to identify explicit pedagogical claims. Objectives were categorized as (a) test-taking strategies, (b) listening skill development, (c) communicative competence development (based on Bachman & Palmer's model), or (d) mixed/unclear aims. Intercoder reliability was established through independent coding of 20% of the data (Cohen's Kappa = 0.84), resolving discrepancies through discussion and coding manual refinement. Direct quotes were extracted to ensure accurate interpretation and resolved mixed cases through collaborative review.

Alignment assessment: Task distributions were evaluated against each coursebook's stated objectives, focusing on communicative authenticity as defined by Bachman and Palmer's (1996) model. This model identifies five core components of language competence: grammatical, textual, functional, sociolinguistic, and strategic. Goh's task types were mapped onto these components to assess communicative demands. One-way transactional tasks mainly target grammatical and textual competence, offering limited authenticity. One-way interactional tasks incorporate sociolinguistic elements, offering moderate authenticity. Two-way transactional tasks require functional and strategic competence, indicating higher authenticity. Two-way interactional tasks integrate functional, sociolinguistic, and strategic dimensions, representing the highest level of communicative authenticity. While all task types involve some grammatical processing, those demanding pragmatic interpretation and adaptive strategies were seen as better aligned with real-world communication and more supportive of communicative competence development.

The alignment analysis examined convergent patterns (where task distributions supported stated aims) and divergent patterns (where content exceeded or fell short of objectives). Of particular interest was positive divergence, where coursebooks included substantially more communicatively authentic tasks than their test-focused statements would predict.

Implications analysis: Finally, the study examined how alignment patterns might impact learner preparation, particularly where IELTS functions as a high-stakes gatekeeping tool. This analysis considered potential consequences of misalignment between stated objectives and actual task provision for learners' communicative development and test performance.

Reliability and validity

Inter-rater reliability

Two trained raters classified all 565 listening tasks using a structured protocol: orientation (4 hours), guided practice (20 tasks, 4 hours), and calibration (15 tasks, 82% agreement). Initial agreement for the full dataset was 79%, with Cohen's kappa of $\kappa = 0.84$ ($p < 0.001$), indicating strong reliability (Landis & Koch, 1977). All discrepancies were resolved through discussion, achieving 100% consensus. To ensure consistency across task types, the classification protocol was validated through expert review, confirming its applicability to diverse task formats, including those with multimedia elements.

Construct validity

Content validity was established through a four-member expert panel review, confirming comprehensive coverage of authenticity features aligned with Bachman and Palmer's (1996) framework. For convergent validity, three experienced IELTS instructors independently rated 100 tasks (17.7% of the dataset) for real-world authenticity on a 5-point scale. Tasks were distributed approximately equally across Goh's (2001) typology categories: two-way interactional ($n = 26$), two-way transactional ($n = 24$), one-way interactional ($n = 25$), and one-way transactional ($n = 25$). Inter-rater reliability for these ratings was high (intraclass correlation coefficient, ICC = 0.87), ensuring consistency. Mean ratings varied by category: two-way interactional ($M = 4.2$), two-way transactional ($M = 3.8$), one-way interactional ($M = 3.1$), and one-way transactional ($M = 2.3$). A one-way ANOVA confirmed significant differences across categories ($F(3,96) = 42.3$, $p < .001$, $\eta^2 = .57$), supporting strong construct validity. The lower authenticity ratings for one-way transactional tasks, often less interactive (e.g., lectures), align with their simpler classification criteria, explaining their high reliability in task categorization despite reduced communicative realism.

Ethical considerations

This study analyzed published IELTS preparation coursebooks without human participants, requiring no IRB approval. The research adhered to academic ethical standards through properly cited excerpts for illustrative purposes, ensuring fair use compliance. Language institutes provided voluntary, non-sensitive input to identify widely used coursebooks based on informed consent. Researchers declare no conflicts of interest or publisher affiliations, ensuring impartial analysis.

Results

Distribution of Listening Task Types (RQ1)

In response to RQ1, the analysis of 565 listening tasks from six IELTS preparation coursebooks showed a heavily skewed distribution toward one-way transactional activities. As displayed in Table 2, one-way transactional tasks made up 70.3% (n = 397) of all activities, which indicated a strong emphasis on non-reciprocal, information-focused listening practice.

Table 2

Distribution of Tasks by Type

Task Type	Frequency	Percentage	95% CI
One-way Transactional	397	70.3%	66.5%–74.0%
One-way Interactional	65	11.5%	8.9%–14.1%
Two-way Transactional	33	5.8%	3.9%–7.8%
Two-way Interactional	70	12.4%	9.7%–15.1%
Total	**565**	**100%**	

The remaining task types were considerably underrepresented: one-way interactional tasks made up 11.5% (n = 65), two-way interactional tasks 12.4% (n = 70), and two-way transactional tasks only 5.8% (n = 33). This distribution highlights a clear imbalance toward monologic, factual listening over interactive or socially focused practice.

Chi-square analysis revealed statistically significant differences in task-type distribution among coursebooks ($\chi^2 = 141.5$, df = 15, $p < 0.001$), though nearly all materials were dominated by one-way transactional formats. The effect size (Cramér's V = 0.29) indicated a moderate association between coursebook selection and task-type emphasis.

Activity-Type Distribution

Table 3 presents the five most frequently occurring listening activity types across all analyzed coursebooks. These formats collectively account for over half of the total tasks and reflect a strong emphasis on traditional test-oriented practice.

Table 3

Top Five Listening Activity Types

Activity Type	Count	Percentage
Multiple-choice	77	13.6%
Matching	66	11.6%
Fill in the Gap	61	10.8%
Pronunciation	57	10.1%
Short-Answer Questions	39	6.9%
Total	**300 out of 565**	**53%**

The top five activity types, mostly focused on assessment and comprehension, make up 53% of all 565 tasks, indicating a clear preference for test-oriented formats. Interactive tasks, such as discussion-based listening, were rare at just 2.6%.

Prevalence of Interactive Tasks (RQ2)

Regarding RQ2, the analysis revealed a clear lack of tasks that support negotiation of meaning and pragmatic understanding. Only 18.2% (n = 103) of all listening tasks involved two-way communication, including two-way transactional tasks (n = 33, 5.8%) and interactional tasks (n = 70, 12.4%). In other words, more than four-fifths of the tasks (81.8%) required no listener interaction, resulting in a predominantly passive listening experience.

From a communicative function standpoint, just 23.9% (n = 135) of tasks had interactional goals (One-way Interactional n=65, 11.5% and Two-way Interactional n=70, 12.4%), while the majority (76.1%, n = 430) were transactional. The low share of two-way interactional tasks (12.4%) is especially striking given their value for developing real-world listening skills.

Individual Coursebooks

Mindset for IELTS: The Mindset series demonstrated a progressive increase in communicative authenticity across proficiency levels, as shown in Table 4. This trend is visually summarized in Figure 1.

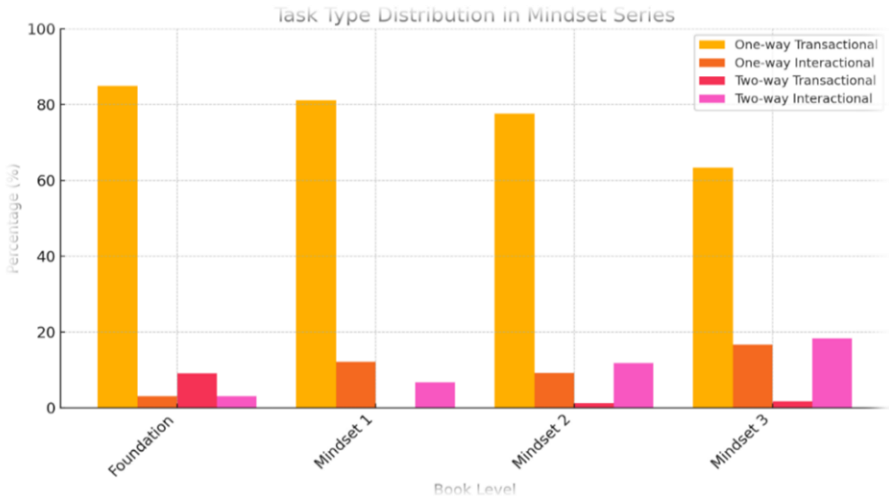
Table 4

Mindset for IELTS Task Types by Proficiency Level

Level	One-way Transactional	One-way Interactional	Two-way Transactional	Two-way Interactional	Total
Foundation	28 (84.8%)	1 (3.1%)	3 (9.1%)	1 (3.0%)	**33** (100%)
Level 1	60 (81.1%)	9 (12.2%)	0 (0.0%)	5 (6.7%)	**74** (100%)
Level 2	59 (77.6%)	7 (9.2%)	1 (1.3%)	9(11.8%)	**76** (100%)
Level 3	38 (63.3%)	10(16.7%)	1 (1.7%)	11 (18.3%)	**60** (100%)
Total	**185 (76.1%)**	**27 (11.1%)**	**5 (2.0%)**	**26 (10.6%)**	**243**(100%)

Figure 1

Distribution of Listening Task Types across the Mindset for the IELTS Series



Chi-square analysis showed significant differences across proficiency levels ($\chi^2 = 21.76$, df = 9, $p < 0.01$, Cramér's V = 0.17), indicating a deliberate progression toward more communicative authenticity at higher levels. The share of two-way interactional tasks rose from 3.0% at the Foundation level to 18.3% at Level 3, a six-fold increase in interactive listening practice.

Complete IELTS: The Complete IELTS series showed a descriptive increase in interactive tasks as proficiency levels rose as presented in Table 5. Figure 2 illustrates the distribution of these task types across the three proficiency bands.

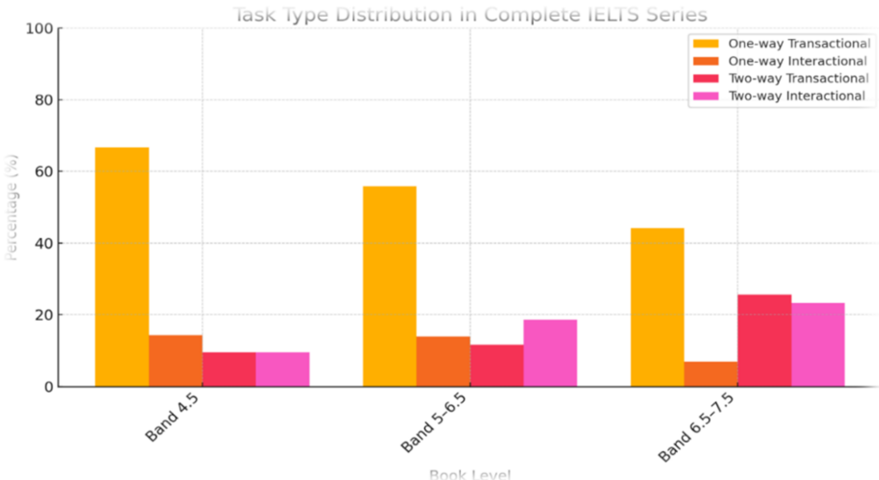
Table 5

Complete IELTS Task Types by Proficiency Levels

Band	One-way Transactional	One-way Interactional	Two-way Transactional	Two-way Interactional	Total
Band 4.5	42 (66.7%)	9 (14.3%)	6 (9.5%)	6(9.5%)	**63**(100%)
Band 5-6.5	24 (55.8%)	6 (14.0%)	5 (11.6%)	8(18.6%)	**43** (100%)
Band 6.5-7.5	19 (44.2%)	3 (7.0%)	11 (25.6%)	10 (23.3%)	**43** (100%)
Total	**85** (57%)	**18** (12%)	**22** (14.7)	**24** (16.1%)	**149** (100%)

Figure 2

Distribution of Listening Task Types across the Complete IELTS Series



The combined share of two-way tasks (transactional + interactional) grew from 19.0% at Band 4.5 to 48.9% at Band 6.5–7.5. Although the chi-square test suggested a shift toward greater communicative authenticity at higher levels (χ^2 = 11.68, df = 6, p = 0.069, Cramér's V = 0.16), this change was not statistically significant. Despite showing the largest descriptive increase among the materials analyzed, the non-significant p-value ($p > 0.05$) means the progression cannot be considered systematic or generalizable beyond this sample. This contrasts with the significant progression observed in the Mindset series, where changes across levels were statistically significant.

The other four coursebooks displayed considerable differences in their communicative orientations:

Collins Listening for IELTS: Exhibited the most restrictive approach, with 90.4% one-way transactional tasks and no two-way activities, representing the most test-oriented profile in the corpus.

Table 6

Collins Listening for IELTS Task Types

Task Type	Count	Percentage
One-Way Transactional	75	90.4 %
One-Way Interactional	8	9.6 %
Two-Way Transactional	0	0 %
Two-Way Interactional	0	0 %
Total	**83**	100

Ready for IELTS: Demonstrated the highest proportion of two-way interactional tasks (41.7%), indicating the strongest commitment to communicative authenticity among individual coursebooks.

Table 7

Ready for IELTS Task Types

Task Types	Count	Percentage
One-way Transactional	9	37.5%
One-way Interactional	2	8.3%
Two-way Transactional	3	12.5%
Two-way Interactional	10	41.7%
Total	**24**	**100.0%**

Ready for IELTS had the highest proportion of two-way interactional tasks (41.7%), despite minimal focus on communicative orientation in its introduction. This suggests that stated teaching philosophy does not always align with task authenticity. In contrast, the Mindset series, though known for emphasizing real-world preparation, showed no such alignment, revealing a gap between its claims and actual content.

IELTS Advantage: It demonstrated a strong communicative focus, with 31.3% two-way interactional and 28.1% one-way interactional tasks, adding up to 59.4% interactional practice overall.

Table 8

IELTS Advantage Task Types

Combined Category	Count	Percentage
One-way transactional	13	40.62%
One-way Interactional	9	28.1%
Two-way Interactional	10	31.3%
Two-way Transactional	0	0.0%
Total	32	100.0%

Barron's IELTS: It maintained a mainly transactional focus, with 88.2% one-way transactional tasks and only 2.9% one-way interactional tasks, reflecting a strong exam-oriented approach.

Table 9

Barron's Task Types

Combined Category	Count	Percentage
One-way Transactional	30	88.2%
One-way Interactional	1	2.9%
Two-way Transactional	3	8.8%
Two-way Interactional	0	0.0%
Total	34	100.0%

Comparative Analysis

Chi-square analysis showed a significant difference in communicative authenticity between *Mindset for IELTS* and *Complete IELTS* ($\chi^2 = 28.74$, df = 3, $p < 0.001$), with a moderate-to-strong effect size (Cramér's V = 0.27). This indicates a strong link between coursebook choice and task emphasis. *Complete IELTS* had more interactional tasks, especially at advanced levels, while *Mindset* showed a gradual increase in complexity. The divergence highlights *Complete IELTS*'s stronger focus on authentic listening, despite both claiming real-world alignment, which is important for curricula valuing interactive skills.

In contrast, *Ready for IELTS* and *IELTS Advantage* showed no significant difference ($\chi^2 = 2.98$, df = 1, p = 0.084) despite *Ready for IELTS* having more two-way tasks (54.2% vs. 31.3%), likely due to the smaller sample size (total n = 56). Both emphasize interactional practice, but *Ready for IELTS* includes more transactional tasks, while *IELTS Advantage* lacks two-way transactional tasks.

Collins and *Barron's* were excluded due to very low two-way tasks (0% and 8.8%) and their heavy one-way transactional focus (Collins 83.1%, Barron's 88.2%), which would skew comparisons.

Mindset and *Complete IELTS* were not compared with *Ready for IELTS* or *IELTS Advantage* because the former are multi-level series with progression, while the latter are standalone volumes, making comparisons invalid.

Alignment with Stated Objectives (RQ3)

Addressing RQ3, this study utilized Bachman and Palmer's (1996) communication model and task analysis to evaluate the correspondence between stated pedagogical goals and actual task design. The findings indicate a prevalent misalignment between stated pedagogical objectives and actual task design, though the nature of this misalignment varies. While most coursebooks (such as *Mindset* and *Collins*) claim to foster communicative competence yet rely heavily on test-focused drills (negative divergence), *Ready for IELTS* demonstrated positive divergence by offering significantly more interactive practice than its exam-focused introduction promised. The detailed alignment analysis for each coursebook follows below:

Ready for IELTS:

The *Ready for IELTS* coursebook presents itself as primarily exam-focused, emphasizing systematic test preparation and general language improvement: "*All four parts of the exam are systematically developed and practiced*," and "*The book contains a wide range of activities aimed at improving your English and developing your band score*." It makes no explicit reference to communicative competence or authentic listening.

Task analysis shows strong communicative authenticity, with 54.2% of listening activities classified as two-way transactional or interactional, indicating implicit alignment with CLT principles. These tasks develop functional competence (goal-oriented negotiation) and strategic competence (clarification, turn-taking), going beyond the coursebook's exam-focused aims. This positive divergence suggests *Ready for IELTS* also supports pragmatic and interactive listening skills.

- Mindset for IELTS (Foundation to level 3)

The *Mindset for IELTS* series emphasizes exam readiness, presenting listening primarily as a receptive, test-oriented skill. Its introduction frames the content as structured preparation for test day: "*MINDSET immerses you in a wide range of IELTS topics and guides you clearly through all the skills and strategies you need… Familiarize yourself with the IELTS test through authentic tasks*."

However, only 12.6% (n = 31) of listening tasks across all levels are communicatively authentic (comprising 2.0% two-way transactional and 10.6% two-way interactional). While this proportion rises to 20.0% at Level 3 (n = 12), the vast majority of the series remains focused on grammatical and textual competence, evidenced by the 76.1% dominance of one-way transactional tasks. Mapped onto Bachman and Palmer's framework, support for sociolinguistic competence is limited (11.1%), and strategic competence, represented by two-way transactional tasks, is negligible at just 2.0%. This distribution reflects a clear gap between the series's claims of authenticity and the actual communicative demands of its tasks.

- *IELTS Advantage: Speaking and listening skills*

IELTS Advantage: Speaking and Listening Skills explicitly adopts an exam-focused approach, emphasizing format familiarity and strategy training*: "The content of the listening section has three main features: format explanation, exam tips… and development of strategies for listening and extracting the information necessary to answer the questions."*

Despite this, task analysis reveals notable communicative authenticity, with 59.4% interactional tasks, 31.3% two-way, and 28.1% one-way. According to Bachman and Palmer's framework, the coursebook shows strong sociolinguistic competence, especially in inferring speaker attitudes and context. However, the absence of two-way transactional tasks (0%) reveals a gap in functional competence, limiting preparation for real-world communication despite a more balanced interactional focus than other test-prep books.

- Barron's IELTS

The introductory description of *Barron's IELTS* states: *"IELTS consists of four sections that test the full range of English language skills… These four modules are examined in detail in this book."*

Though covering all skills, the listening section is mainly exam-focused, with 88.2% one-way transactional tasks and no two-way interactional tasks. Two-way transactional tasks make up only 8.8%, and one-way interactional tasks just 2.9%, providing limited opportunities for real-world communication. Compared to *IELTS Advantage* (59.3% authentic tasks) and *Complete IELTS* (30.9%), *Barron's* shows the weakest alignment with communicative teaching principles.

Using Bachman and Palmer's framework, the coursebook focuses on textual competence and detail extraction. Sociolinguistic competence is minimal (2.9%), and strategic competence is absent. While this may help with test familiarity, it does not adequately develop the functional listening skills needed for academic and professional success.

- Collins Listening for IELTS

Collins Listening for IELTS presents itself as a test-focused resource: *"Units 1–11 cover the key types of questions that you find in the IELTS listening test. Every exercise is relevant to the exam."*

This focus shows in its task distribution: 83.1% one-way transactional, with no two-way interactional or transactional tasks. Only 9.6% are one-way interactional, limiting attention to inference and speaker intent. The absence of reciprocal interaction greatly restricts communicative competence development.

According to Bachman and Palmer's framework, the coursebook emphasizes grammatical and textual competence but neglects pragmatic and strategic skills. This narrow focus on test mechanics requires instructors to provide substantial supplementary materials to develop interactive and strategic listening skills.

- Complete IELTS

The *Complete IELTS* series positions itself as an integrated skill development resource, stating: "*It teaches you the reading, writing, listening, and speaking skills that you need for the exam. It covers all the exam question types, as well as key grammar and vocabulary, which, from research into the Cambridge Learner Corpus, are known to be useful to candidates needing to achieve a high band score in the test*."

Its exam-focused approach includes 57.1% one-way transactional tasks, with 16.1% two-way interactional and 14.8% two-way transactional tasks, totalling 30.9% communicatively authentic activities. This reflects a moderate balance, with some interaction-rich tasks supporting test preparation.

Mapped to Bachman and Palmer's framework, the coursebook demonstrates moderate functional and strategic competence, especially at higher levels, with 48.9% two-way tasks at Bands 6.5 to 7.5. Although not highlighted in the description, authentic tasks suggest underlying support for communicative competence that teachers can build on with extra speaking and interaction activities. Overall, *Complete IELTS* partly meets its integrated skills claim by balancing exam preparation with real-world communication opportunities.

Overall, IELTS preparation materials tend to offer mostly passive, test-focused listening practice, providing limited chances for interactive, pragmatically rich experiences essential to authentic communicative competence.

Discussion

Dominance of One-way Transactional Tasks

The finding that 70.3% of the 565 analyzed listening tasks in IELTS preparation coursebooks are one-way transactional underscores a significant tension between test-oriented pedagogy and the principles of Communicative Language Teaching. This predominance, as shown in Table 2, emphasizes passive information extraction, aligning with previous research highlighting the prevalence of literal comprehension tasks in IELTS preparation contexts (Baghaei et al., 2020; Field, 2012), though the extent suggests a more evident instructional narrowing than previously documented.

From a psycholinguistic viewpoint, this distribution does not engage the cognitive processes involved in authentic listening (Rost & Brown, 2022). The dominance of one-way transactional tasks reduces listening to recognition exercises, neglecting the executive functions, predictive processing, and social cognition that are important for real-world communication. EFL learners, who rely heavily on these materials in contexts and have limited exposure to authentic English (Babaee, 2017), may develop listening skills tailored to tests, but they become poorly prepared for interactive academic and professional environments. This concern is supported by research suggesting that intensive test preparation can cause considerable harm to performance on authentic listening tasks (Jalilzadeh et al., 2023; Wen et al., 2022), indicating possible negative washback effects from current teaching practices.

Scarcity of Interactive Tasks

The finding that only 18.2% of listening tasks involve two-way communication shows a significant gap in IELTS preparation materials. This shortage is concerning, given that negotiation of meaning plays a central role in second language acquisition theory (Long, 1996) and communicative language teaching practice (Thornbury, 2016). Interactive tasks give learners chances to clarify, confirm, and build understanding together, processes that are essential for developing the pragmatic skills needed in real-world communication.

The near absence of two-way transactional tasks (5.8%) is particularly important, as these activities help bridge the gap between test-focused practice and authentic communication by requiring learners to negotiate meaning toward specific goals while still focusing on information exchange. This finding highlights the importance of orienting EFL learners toward communicative acts through task planning (Ballesteros Soria & van Compernolle, 2023).

Variation among Coursebooks

The large variation in communicative authenticity among coursebooks (ranging from 0% to 54.2% authentic tasks) shows important tensions within the IELTS preparation market. The different approaches seen in Collins Listening for IELTS (90.4% one-way transactional) and Ready for IELTS (41.7% two-way interactional) reflect two distinct teaching approaches: narrow test preparation versus broader language development.

This imbalance likely reflects three market realities: test-focused materials are seen as safer investments, one-way transactional tasks are cheaper to produce, and publishers worry that innovative materials might not be seen as serious test preparation. This creates a cycle where publishers produce what sells, institutions buy what looks like 'real' test preparation, and learners get stuck with approaches that focus on form over function.

The Cambridge series (Mindset and Complete IELTS) takes a middle position, showing systematic progression toward greater communicative authenticity at higher proficiency levels. This suggests that publishers can be aware of learner needs across different proficiency levels.

Misalignment

The consistent pattern of differences between stated teaching goals and actual task distributions shows a significant gap between theory and practice. This mismatch may reflect authors' lack of awareness about how task choices align with communicative principles, market pressures favoring test-focused content, or uncertainty about balancing exam preparation with communicative development.

Conclusion

This study highlights a clear gap between the listening tasks commonly found in popular IELTS preparation coursebooks and the realities of authentic communication. Although the books differ in quality, some being highly exam-focused and others offering more interactive elements, the overall pattern is the dominance of one-way, passive information-retrieval tasks. In EFL contexts such as Iran, where learners have limited exposure to English outside the classroom, this overreliance on monologic tasks can restrict learners' development. It shows that, despite claims of promoting communicative competence, many commercial materials unintentionally encourage negative washback, prioritizing test scores over the pragmatic and strategic listening skills needed for academic and professional settings.

The findings have important implications for three main groups: teachers, material developers, and policymakers. For teachers, the lack of two-way, interactional tasks means they cannot rely solely on coursebooks. They need to build in their own activities that encourage negotiation of meaning, such as prediction tasks completed in pairs or post-listening information-gap activities. Classroom assessment practices also need to shift toward valuing interactional skills, like turn-taking and asking for clarification, rather than focusing only on getting the right answer. This reframes the during-listening stage as a space for discussion rather than silence. For material developers, there is a clear need to align content with

Communicative Language Teaching principles by using multimedia, a wider range of accents, and tasks that require back-and-forth communication. Policymakers, in turn, must recognize that communicative curriculum reforms will remain largely symbolic unless accompanied by materials that genuinely balance test preparation with skill development.

While this study provides an in-depth review of six widely used coursebooks, it focuses only on printed materials and may not represent the full range of IELTS resources available globally. Future research should include digital platforms and mobile applications, which now play a major role in learner preparation. Investigating how teachers and learners interpret and use these tasks in real classrooms could also uncover practical issues that a text-based analysis cannot reveal. In addition, longitudinal studies tracking how exposure to communicative tasks affects test performance would offer stronger evidence for the improvements recommended here.

Acknowledgments

The authors thank the applied linguistics experts and IELTS instructors who contributed to the expert review and pilot testing of the framework adaptation, whose input ensured the reliability and validity of the classification procedures, as well as the directors of the language institutes who voluntarily shared information about coursebook adoption patterns in Iran.

References

Aish, F., & Tomlinson, J. (2019). *Collins listening for IELTS* (Revised ed.). HarperCollins Publishers.

Alavi, S. M., Kaivanpanah, S., & Masjedlou, A. P. (2018). Validity of the listening module of international English language testing system: Multiple sources of evidence. *Language Testing in Asia*, *8*(1), 1-17. https://doi.org/10.1186/s40468-018-0057-4

Archer, G., Kosta, J., & Passmore, L. (2018). *Mindset for IELTS: Level 3*. Cambridge University Press.

Archer, G., Kosta, J., Passmore, L., & Uddin, J. (2017). *Mindset for IELTS: Foundation*. Cambridge University Press.

Aryadoust, V. (2011). Validity arguments of the speaking and listening modules of international English language testing system: a synthesis of existing research. *The Asian ESP Journal, 7*(1), 28–54.

Babaee, H. (2017). A comprehensive look into the instruction of listening skill in academic English programs: A case study of two state universities in Iran. *International Journal of Education and Literacy Studies*, *5*(2), 112–125. https://doi.org/10.7575/aiac.ijels.v.5n.2p.112

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

Badpa, H., & Beikian, A. (2025). Evaluating overall quality and intercultural communicative competence in English academic writing textbooks: A comparative study of Iranian and international resources. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, *13*(2), 131-154. https://doi.org/10.22049/jalda.2025.30314.1761

Baghaei, S., Bagheri, M. S., & Yamini, M. (2020). Analysis of IELTS and TOEFL reading and listening tests in terms of Revised Bloom's taxonomy. *Cogent Education, 7*(1), Article 1720939. https://doi.org/10.1080/2331186X.2020.1720939

Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, *13*(3), 257-279. https://doi.org/10.1177/026553229601300303

Ballesteros Soria, N., & van Compernolle, R. A. (2023). Collectivizing an orientation to turn-allocation as a learnable through pre-task planning. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, *11*(2), 133-160. https://doi.org/10.22049/jalda.2024.28178.1502

Bouftira, M., El Messaoudi, M., & Li, S. (2022). Developing EFL learners' pragmatic competence through a blended learning model: A quasi-experimental study. *European Scientific Journal, 18*(16), 105–124. https://doi.org/10.19044/esj.2022.v18n16p105

Brook-Hart, G., & Jakeman, V. (2012). *Complete IELTS Bands 5-6.5 Student's Book*. Cambridge University Press.

Brook-Hart, G., & Jakeman, V. (2013). *Complete IELTS Bands 6.5–7.5 Student's Book*. Cambridge University Press.

Brown, G., & Yule, G. (1983). *Teaching the spoken language*. Cambridge University Press.

Brunfaut, T., & Révész, A. (2015). The role of task and listener characteristics in second language listening. *TESOL Quarterly, 49*(1), 141–168. https://doi.org/10.1002/tesq.168

Buck, G. (2001). *Assessing listening*. Cambridge University Press. https://doi.org/10.1017/CBO9780511732959

Chalmers, J., & Walkinshaw, I. (2014). Reading strategies in IELTS tests: Prevalence and impact on outcomes. *English Australia Journal, 30*(1), 24–39.

Cheng, L., Sun, Y., & Ma, J. (2015). Review of washback research literature within Kane's argument-based validation framework. *Language Teaching, 48*(4), 436–470. https://doi.org/10.1017/S0261444815000233

Cheng, L., Watanabe, Y., & Curtis, W. A. (Eds.). (2004). *Washback in language testing: Research contexts and methods*. Routledge. https://doi.org/10.4324/9781410609731

Crosthwaite, P., Hutchison, S., Wijayatilake, C., & De Souza, N. (2017). *Mindset for IELTS: Level 1*. Cambridge University Press.

De Souza, N. (2017). *Mindset for IELTS: Level 2*. Cambridge University Press.

Esmaeili, S. Z., Kuhi, D., & Behroozizad, S. (2022). Cultivating intercultural communicative competence through interculturally-laden tasks: A qualitative inquiry. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, *10*(2), 183-209. https://doi.org/10.22049/jalda.2022.27642.1389

Estaji, M., & Ghiasvand, F. (2021). IELTS washback effect and instructional planning: The role of IELTS-related experiences of Iranian EFL teachers. *Journal of English Language Teaching and Learning*, *13*(27), 163-192. https://doi.org/10.22034/elt.2021.45369.2369

Field, J. (2003). Assessing listening. *ELT Journal*, *57*(2), 198-202. https://doi.org/10.1093/elt/57.2.198-a

Field, J. (2012). The cognitive validity of the lecture-based question in the IELTS listening paper. *IELTS collected papers*, *2*, 391-453.

Field, J. (2023). *Insights into assessing academic listening: The case of IELTS*. CRELLA, University of Bedfordshire.

Gilabert, R., Barón, J., & Llanes, À. (2009). Manipulating cognitive complexity across task types and its impact on learners' interaction during oral performance. *International Review of Applied Linguistics in Language Teaching, 47*(3–4), 367–395. https://doi.org/10.1515/iral.2009.016

Gilmore, A. (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching*, *40*(2), 97-118. https://doi.org/10.1017/S0261444807004144

Goh, C. C. M. (2000). A cognitive perspective on language learners' listening comprehension problems. *System, 28*(1), 55–75. https://doi.org/10.1016/S0346-251X(99)00060-3

Goh, C. C. M. (2001). Designing listening comprehension tasks. *Guidelines, 23*(2), 4-11

Goh, C. C. M., & Vandergrift, L. (2012). *Teaching and learning second language listening: Metacognition in action*. Routledge. https://doi.org/10.4324/9780203843376

Green, A. (2007). *IELTS washback in context: Preparation for academic writing in higher education*. Cambridge University Press.

Hasrol, S. B., Zakaria, A., & Aryadoust, V. (2022). A systematic review of authenticity in second language assessment. *Research Methods in Applied Linguistics, 1*(3), Article 100023. https://doi.org/10.1016/j.rmal.2022.100023

Huang, L., & Zeng, J. (2025). Sustainability of washback effects: A longitudinal study of China's twice-yearly NMET reform. *Language Testing in Asia, 15*, 19. https://doi.org/10.1186/s40468-025-00358-9

Hughes, A. (1989). *Testing for language teachers*. Cambridge University Press. https://doi.org/10.1017/CBO9780511732980

Jalilzadeh, K., Dastgoshadeh, A., & Khosravi, R. (2023). Language assessment practices and training preferences of EL teachers: Iranian EFL teachers in focus. *Applied Research on English Language*, *12*(1), 27-54. https://doi.org/10.22034/elt.2023.59200.2583

Kirkpatrick, A. (2020). Englishes in the expanding circle: Focus on Asia. *Russian Journal of Linguistics, 24*(3), 551–568. https://doi.org/10.22363/2687-0088-2020-24-3-551-568

Knoch, U., Huisman, A., Elder, C., Kong, X., & McKenna, A. (2020). *Drawing on repeat test takers to study test preparation practices and their links to score gains*. Language Testing. https://doi.org/10.1177/0265532220927407

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

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.

Lougheed, L. (2016). *Barron's IELTS* (4th ed.). Barron's Educational Series.

Marks, J. (2017). *IELTS advantage: Speaking and listening skills*. Delta Publishing by Klett.

Marina, K. (2018). The validation process in the IELTS reading component: Reading requirements for preparing international students. *Journal of Language and Education*, *4*(1), 63-78. https://doi.org/10.17323/2411-7390-2018-4-1-63-78

McCarter, S. (2017). *Ready for IELTS* (2nd ed.). Macmillan Education

Mickan, P., & Motteram, J. (2009). The preparation practices of IELTS candidates: Case studies. *IELTS Research Reports, 10*, 223–262.

Minh, N. T. (2025). IELTS washback as a high-stakes test on student learning: A hierarchical modelling study at a Vietnamese university. *International Journal of Technology and Educational Studies, 5*(1), 1–15

Moharami, M., & Daneshfar, S. (2022). The political climate of English language education in Iran: A review of policy responses to cultural hegemony. *Issues in Educational Research, 32*(1), 248–263. https://search.informit.org/doi/10.3316/informit.475768620543197

Nguyen, H. T. M., & Le, N. V. A. (2024). Text complexity of Cambridge-delivered IELTS Academic Reading tests: Comparability with IELTS Academic Reading practice tests from other publishers. *TESL-EJ, 28*(2). https://doi.org/10.55593/ej.28110a4

Nunan, D. (1999). *Second language teaching & learning*. Heinle & Heinle.

Ockey, G. J., & Wagner, E. (2018). *Assessing L2 listening: Moving towards authenticity*. John Benjamins.

Peacock, M. (1997). The effect of authentic materials on the motivation of EFL learners. *ELT Journal*, *51*(2), 144-156. https://doi.org/10.1093/elt/51.2.144

Phakiti, A. (2016). Test-takers' performance appraisals, appraisal calibration, state-trait strategy use, and state-trait IELTS listening difficulty in a simulated IELTS listening test. *IELTS Research Reports Online Series, 6,* 1–3.

Priyanti, N. (2017). Test usefulness of IELTS writing test tasks. *Apples-Journal of Applied Language Studies*, *11*(4), 1-9. https://doi.org/10.17011/apples/urn.201708073427

Rasti, A. (2018). Iranian EFL teachers' sense-making of policy reforms: The case of the new communicative-based curriculum. *Journal of Teaching Language Skills*, *37*(3), 169–193. https://doi.org/10.22099/jtls.2019.30819.2580

Read, J. (2022). Test review: The international English language testing system (IELTS). *Language Testing*, *39*(4), 679-694. https://doi.org/10.1177/02655322221086211

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, *22*(1), 27-57. https://doi.org/10.1093/applin/22.1.27

Rost, M., & Brown, S. (2022). Second language listening. In M. Griffiths (Ed.), *Handbook of Practical Second Language Teaching and Learning* (pp. 233-250). Routledge. https://doi.org/10.4324/9781003106609-20

Saif, S. (2021). High-stakes test preparation in Iran: The interplay of pedagogy, test content, and context. In B. Lanteigne, C. Coombe, & J. D. Brown (Eds.), *Challenges in Language Testing Around the World* (pp. 777-799). Springer. https://doi.org/10.1007/978-981-33-4232-3_35

Tajeddin, Z., & Chamani, F. (2020). Foreign language education policy (FLEP) in Iran: Unpacking state mandates in major national policy documents. *Journal of Teaching Language Skills*, *39*(39), 185–215. https://doi.org/10.22099/jtls.2021.38870.2904

Thornbury, S. (2016). Communicative language teaching in theory and practice. In *The Routledge handbook of English language teaching* (pp. 224-237). Routledge. https://doi.org/10.4324/9781315676203-20

Tsagari, D., & Cheng, L. (2017). Washback, impact, and consequences revisited. In E. Shohamy, I. Or, & S. May (Eds.), *Language Testing and Assessment* (pp. 359-372). Springer. https://doi.org/10.1007/978-3-319-02261-1_24

Wen, Z., Zhou, Y., Li, M., Lu, Y., & Chen, Y. (2022). The washback of IELTS listening on senior school English listening teaching and learning. *International Journal of Education and Humanities*, *4*(3), 20–27. https://doi.org/10.54097/ijeh.v4i3.1646

Wyatt, R. (2012). *Complete IELTS bands 4-5 student's book*. Cambridge University Press.

Authors' Biographies

Sasan Baleghizadeh is Professor of TEFL at Shahid Beheshti University in Tehran, Iran, where he teaches courses in applied linguistics, syllabus design, and materials development. He is interested in investigating the role of interaction in English language teaching and issues related to materials development. His published articles appear in both national and international journals including *TESL Reporter*, *ELT Journal*, *Language Learning Journal*, and *Language Teaching Research*.

Mehran Mousavi is a PhD student of TEFL at Shahid Beheshti University in Tehran, Iran. His main areas of interest include academic writing, language assessment, and IELTS preparation.

Volume 14, Issue 1
Winter and Spring, 2026
pp. 147-176

Journal of Applied Linguistics and Applied Literature: Dynamics and Advances

Reading with Support: The Influence of Scaffolding Models on EFL Learners' Academic Selves, Resilience, and Self-Regulation

Komeil Jafari Sanksarak[1], Mehrshad Ahmadian[2*] and Ebrahim Fakhri Alamdari[3]

[1]*PhD Candidate in ELT, QaS.C., Islamic Azad University, Qaemshahr, Iran ORCID: 0009-0001-7737-3046;* komeil.jafari5227@iau.ac.ir

[2]*Corresponding Author: Assistant Professor of TEFL, Department of Foreign Languages, QaS.C., Islamic Azad University, Qaemshahr, Iran ORCID: 0000-0002-2105-6818;* mehrshad.ahmadian@iau.ac.ir

[3]*Assistant Professor of TEFL, Department of Foreign Languages, QaS.C., Islamic Azad University, Qaemshahr, Iran, ORCID: 0000-0001-5075-3703* ebrahim.fakhri@qaemiau.ac.ir

Abstract

This mixed-methods study examined the effects of three scaffolded academic reading comprehension instructional models— Joyce-Gibbons' (2017) socioculturally informed classroom-based scaffolding, Van de Pol et al.'s (2010) contingent scaffolding framework, and Coyle's (2015) scaffolding approach on both reading achievement and critical learner characteristics —on EFL undergraduate learners' reading achievement, motivation, resilience, and self-regulation at Islamic Azad University of Mazandaran. A total of 120 Persian-speaking students, aged 19 to 25, from diverse academic majors were randomly assigned to three experimental groups. Standardized instruments including the Oxford Placement Test, L2 Motivational Selves Questionnaire, Resilience Questionnaire, Self-Regulation Questionnaire, and a Reading Comprehension Test were administered pre- and post-intervention. Over 8–10 weeks, each group received scaffolded instruction tailored to its assigned model. Quantitative analyses (MANOVA) revealed significant improvements in all measured domains for each instructional model compared to control groups, with large effect sizes confirming substantial gains in both cognitive and affective learner outcomes. Each of these models, although distinct in their pedagogical underpinnings, shares a common commitment to providing structured, responsive, and gradual support that facilitates learners' transition from assisted to independent performance. Qualitative data from semi-structured interviews further highlighted learners' positive perceptions of scaffolded instruction and its role in enhancing academic and emotional engagement. The findings underscore the pedagogical value of scaffolded reading strategies in fostering holistic learner development and suggest directions for integrating culturally responsive scaffolding approaches in EFL settings.

Keywords: academic achievement, motivation, resilience, scaffolded instruction, self-regulation

ARTICLE INFO
Research Article
Received: Tuesday, August 5, 2025
Accepted: Thursday, February 5, 2026
Published: Wedensday, April 1, 2026
Available Online: Thursday, February 5, 2026
DOI: https://doi.org/ 10.22049/jalda.2026.30880.1850
Online ISSN: 2821-0204; Print ISSN: 28208986

Introduction

In the realm of English as a Foreign Language (EFL) education, the development of learners' academic identity and the cultivation of adaptive learning behaviors have become central to promoting long-term academic success (Abdelshaheed, 2023; Marzban et al., 2021). Among the strategies employed to support EFL learners, scaffolding has emerged as a key pedagogical approach that enables students to navigate complex linguistic and cognitive tasks with structured assistance (Taye & Teshome, 2025). Grounded in sociocultural theory, scaffolding is defined as the provisional assistance offered by teachers or more knowledgeable peers, which is gradually withdrawn as learners become more autonomous (Ahmed Abdel-Al Ibrahim et al., 2023). Beyond facilitating language development, this teaching approach significantly influences the formation of learners' academic self-concepts, including their beliefs, aspirations, and sense of capability in academic contexts (Derakhshan & Fathi, 2025); Wang, 2017).

As modern classrooms increasingly emphasize learner autonomy, resilience, and self-regulation, it becomes imperative to explore how different scaffolding models impact these interconnected constructs (Yang et al., 2025). Academic resilience, defined as the capacity to persevere and succeed despite challenges, and self-regulation, encompassing goal-setting, strategy use, and self-monitoring, are particularly critical for EFL learners who often face additional language-related barriers. The extent to which scaffolding techniques foster these attributes can significantly influence learners' motivation, persistence, and overall academic engagement (Hou, 2025; Wang et al., 2024).

The increasing complexity of academic tasks in EFL contexts requires learners to not only develop linguistic competence but also foster internal capacities such as a strong academic self-concept, emotional resilience, and effective self-regulation strategies (Mellati & Valizadeh, 2025; Taye & Teshome, 2025). However, many EFL learners struggle with feelings of inadequacy, low confidence, and anxiety due to the dual challenge of mastering content and language simultaneously. Without appropriate instructional support, these challenges can hinder their academic growth and long-term engagement. Therefore, understanding how pedagogical interventions—particularly scaffolding models—can influence learners' academic selves, resilience, and self-regulation is both timely and essential (Wang et al., 2024). Scaffolding, based on Vygotsky's sociocultural theory, is a recognized teaching strategy designed to help learners progress from what they can accomplish on their own to what they can achieve with support (Zhang, 2024). Although its role in enhancing language proficiency is well documented, its wider effects on psychological and behavioral aspects within EFL contexts have received comparatively limited attention (Yang et al., 2025; Zhi & Derakhshan, 2024). Specifically, there is a need to examine how different scaffolding models—such as peer-assisted learning, teacher-guided instruction, or digital scaffolding—shape learners' beliefs about their academic capabilities, foster perseverance in the face of difficulty, and promote autonomous learning behaviors (Derakhshan et al., 2024; Li et al., 2025).

Despite the increasing emphasis on student-centered learning and the promotion of higher-order cognitive and affective skills in EFL settings, many university-level language learners continue to struggle with academic reading comprehension and related psychological capacities such as motivation, resilience, and self-regulation (Estrada-Araoz et al., 2023). Traditional instructional approaches often fail to provide the necessary support structures that allow learners to engage deeply with academic texts while simultaneously developing the personal competencies required for sustained academic success. Although scaffolding is widely acknowledged as an effective teaching approach for facilitating language learning, there is still limited understanding of how various scaffolding models differ in their impact on learners' academic performance (Mallahi, 2025). In particular, the specific effects of models such as Joyce-Gibbons' (2017) socioculturally informed classroom-based scaffolding, Van de Pol et al.'s (2010) contingent scaffolding framework, and Coyle's (2015) scaffolding approach on both reading achievement and critical learner characteristics—such as motivational selves, resilience, and self-regulation—remain unclear (Pan et al., 2025).

Therefore, this study seeks to address several interrelated research problems concerning the implementation and impact of scaffolded instruction in EFL academic reading contexts. First, there is an insufficient understanding of how scaffolded instruction based on Joyce-Gibbons (2017) model influences university EFL students' reading comprehension, motivational selves, learning resilience, and self-regulation. Additionally, empirical evidence remains limited regarding the effects of Van de Pol et al. (2010) contingent scaffolding model on the same set of academic and affective outcomes. Similarly, the influence of Coyle (2015) content and language integrated scaffolding approach on students' reading achievement and key psychological learning variables remains unclear. Furthermore, there is a notable gap in knowledge concerning the specific mediation strategies that teachers employ when implementing these different scaffolding models in real instructional settings. Addressing these problems is essential for advancing our understanding of how targeted instructional support can foster both cognitive and affective growth in university-level EFL learners.

Literature Review

The study explored the impact of scaffolded academic reading instruction on EFL learners' motivational selves, learning resilience, self-regulation, and reading achievement, drawing on three prominent scaffolding models. To establish a solid theoretical and empirical foundation for this investigation, the review of the literature examines key constructs and frameworks that inform the study. Specifically, it outlines the concept and evolution of scaffolding in educational contexts, with a focus on its application in language learning and academic reading. The review also highlights major theoretical models of scaffolding, including those proposed by Joyce-Gibbons (2017), Van de Pol et al. (2010), and (Coyle, 2015), emphasizing their pedagogical principles and relevance to EFL instruction. In addition, the literature review investigates three essential learner-related variables—motivational selves, learning resilience, and self-regulation—that play a critical role in academic success, especially in challenging language learning environments. By synthesizing previous studies on these constructs, the review aims to demonstrate how they are influenced by instructional strategies such as scaffolding. Finally, the review addresses the gap in research regarding the mediational strategies used by teachers in scaffolding-rich classrooms and their potential contribution to learner development. Together, these strands of literature provide a comprehensive background for understanding the multifaceted effects of scaffolded instruction and justify the need for the current study.

Scaffolding in Educational Contexts

The concept of scaffolding remains foundational in educational practice, rooted in Vygotsky's sociocultural theory and the notion of the Zone of Proximal Development (ZPD). Recent research continues to highlight its relevance in language education, particularly in fostering learner autonomy and effective comprehension strategies. For example, (Sarmiento-Campos et al., 2022) demonstrated that peer scaffolding, conducted within ZPD-aligned dyads, notably improved collaborative writing and speaking skills among EFL learners. In technology-enhanced EFL contexts, Alfares (2025) investigated intelligent tutoring systems that embedded multimodal scaffolding instructions aligned with ZPD principles, finding that such systems provided adaptive, timely support and boosted language development. Additionally, Ahmed Abdel-Al Ibrahim et al. (2023) employed Functional Magnetic Resonance Imaging (fMRI) to compare scaffolded versus non-scaffolded writing tasks in an EFL setting, discovering that scaffolded instruction led to superior performance and reduced cognitive load in brain regions associated with language processing. These studies collectively reaffirm that scaffolding—whether through peer interaction, digital tutors, or guided task design—continues to play a critical role in breaking down complex linguistic tasks, promoting linguistic growth, and nurturing learner agency and engagement.

Scaffolding Models in Language Education

Three prominent models of scaffolding provide the theoretical backbone of this study. Joyce-Gibbons' (2017) model highlights the importance of dialogic instruction and interactional scaffolding within mainstream educational environments, aiming to merge language and content learning through collaborative classroom dialogue. This approach is especially applicable in EFL settings, where learners gain from repeated, context-rich exposure to academic language forms. In contrast, Van de Pol et al. (2010) introduce a contingent scaffolding framework, in which instructional support is continuously adapted to align with students' evolving needs. This model outlines three essential components: diagnostic strategies to assess understanding, contingency in support based on learner response, and fading as learner competence increases. Its strength lies in the real-time responsiveness of instruction, aligning well with differentiated approaches in EFL classrooms. Coyle (2015) offer a model, which situates language learning within subject-matter instruction. Scaffolding in this model focuses on cognitive challenge, language support, and active learning, fostering academic language proficiency alongside subject understanding. This model is increasingly applied in EFL programs where academic reading is embedded in content-based instruction.

Academic Reading and Scaffolding in EFL Contexts

Reading academic texts in a foreign language poses unique challenges, including unfamiliar vocabulary, complex syntax, and cultural references (Alfares, 2025). EFL learners often require explicit support to decode texts, construct meaning, and develop critical reading strategies (Jeon & Yamashita, 2014). Scaffolded reading instruction has been found to improve comprehension and foster independent reading skills by modeling strategies such as prediction, questioning, summarizing, and clarifying (van de Pol et al., 2021). Moreover, scaffolded approaches encourage engagement with texts through structured interaction, peer collaboration, and teacher-guided discussion, which are vital for developing academic literacy in EFL settings (Teng et al., 2024).

Motivational Selves and Scaffolding

Motivational selves emphasize the role of self-guides in motivating language learners (Grocott et al., 2019; Zhang, Chen & Peng, 2025). Scaffolding can support the development of learners' motivational selves by creating success experiences, increasing self-efficacy, and aligning tasks with personal goals (Li et al., 2025). Research suggests that learners who perceive instructional support as responsive and empowering are more likely to develop a positive academic self-concept and invest effort in learning tasks (You et al., 2016). On the other hand, Motivational selves can affect feedback-seeking behaviour of learners. Mallahi (2025) indicated that ought-to self predicted EFL learners' tope of feedback seeking (Mallahi, 2025)

Learning Resilience in EFL Education

Learning resilience is defined as the capacity to cope with and overcome academic obstacles (Halterman, 2023). This trait is especially critical in EFL environments, where students frequently encounter challenges stemming from language barriers and anxiety related to performance (Wang et al., 2024). Instructional scaffolding can promote resilience by providing structured opportunities to succeed, reducing cognitive overload, and encouraging risk-taking in a supportive environment (Grocott et al., 2019). (Ahmed Abdel-Al Ibrahim et al., 2023).

Self-Regulation and Scaffolding

Self-regulation involves learners' ability to set goals, plan, monitor progress, and reflect on learning outcomes (Baranovskaya, 2015). In scaffolded instruction, teachers' model and gradually transfer these regulatory strategies to learners, helping them become more autonomous. Empirical studies have demonstrated that scaffolded learning environments enhance self-regulatory behaviors by promoting metacognitive awareness and strategic thinking (Danli, 2017). Especially in academic reading, scaffolding strategies such as guided questioning, graphic organizers, and structured summarization help learners manage their reading processes more effectively (Baranovskaya, 2015).

Empirical Studies

A growing body of empirical work has consistently shown that scaffolded instruction plays a meaningful role in shaping EFL learners' academic self-concept, resilience, self-regulation, and reading achievement. Rather than functioning as isolated instructional techniques, scaffolding practices appear to operate as structured forms of support that gradually transfer responsibility from teachers to learners, thereby strengthening learners' sense of competence and control over learning tasks. This process is particularly relevant in academic literacy contexts, where cognitive demands are high and sustained engagement is required.

Recent studies in technology-mediated and classroom-based settings further clarify how scaffolding contributes to these outcomes. For example, Uçak and Kartal (2023) demonstrated that embedding scaffolding within online reading strategy training not only improved learners' comprehension performance but also supported the development of self-regulated learning behaviors. Their findings suggest that scaffolding can compensate for the reduced immediacy of online environments by making strategic processes more explicit and accessible. Similarly, Zhang, Zou and Cheng (2024) reported that explicit instruction in self-regulation strategies led to gains in motivation, self-efficacy, and willingness to communicate, indicating that scaffolding may influence both cognitive processes and affective dimensions of language learning.

Evidence from adjacent skill domains reinforces the broader impact of scaffolded instruction. In academic writing contexts, Allagui (2024) found that scaffolded support in source-based argumentative tasks enhanced learners' writing performance while simultaneously strengthening their self-efficacy beliefs. This dual effect suggests that scaffolding does not merely improve task outcomes but also reshapes learners' perceptions of their own academic capabilities. From a more general educational perspective, Härkki (2020) highlighted the role of adaptive scaffolds in promoting self-regulated learning behaviors, emphasizing the importance of adjusting support in response to learners' evolving needs. This perspective aligns with language-focused research that views scaffolding as a dynamic, responsive process rather than a fixed instructional sequence.

Further support for the interplay between scaffolding, self-regulation, and reading achievement is provided by Mohammadi et al. (2020), who showed that cognitive and metacognitive self-regulation strategies directly predicted EFL learners' reading comprehension, with motivation acting as an indirect mediator. Their findings underline the mechanism through which scaffolded instruction may exert its effects, namely by strengthening learners' strategic engagement and motivational resources. More recent studies have extended these insights by documenting learners' subjective responses to scaffolded instruction. Hassen et al. (2024) reported significant improvements in writing skills following scaffolded instruction, accompanied by positive learner perceptions of clarity, support, and gradual independence. Likewise, Wang and Fan (2025) found that interactive scaffolding increased university EFL learners' engagement and their use of self-regulated reading strategies, highlighting the role of interaction and feedback in sustaining learner involvement.

Halterman (2023) demonstrated that contract-based scaffolding, including goal setting and pacing strategies, helped enhance learners' engagement and persistence, especially among at-risk students. Acosta-Gonzaga and Ramirez-Arellano (2022) learned that scaffolded teacher–student interaction boosted metacognitive strategy awareness and both emotional and behavioral engagement. Finally, Michalsky (2021) examined the timing of motivational scaffolding and found that delivering support before, during, or after reading tasks had significant effects on learners' comprehension and engagement. Collectively, these studies reinforce the efficacy of scaffolded instruction in supporting key psychological and academic variables among EFL learners, forming a strong empirical foundation for investigating the differential effects of various scaffolding models.

Recent empirical studies reveal unresolved issues about how different scaffolding models affect academic and psychological outcomes in EFL learners. Although scaffolding generally benefits academic performance and self-regulation, there is limited agreement on which models are most effective. Gibbons' (2015) model, focused on literacy development, lacks systematic examination of its impact on motivation, resilience, and self-regulation in university reading contexts. Van de Pol et al.'s (2010) contingent scaffolding shows theoretical promise but insufficient long-term evidence on its effects on learner traits like resilience and motivation. Coyle's (2015) CLIL model presents mixed findings about its psychological benefits, especially concerning content complexity and scaffolding adequacy. Additionally, all three models suffer from insufficient documentation of teachers' mediation strategies, hindering understanding of how scaffolding translates into learner outcomes. These gaps motivate the study's research questions, aiming to clarify the comparative effectiveness and mechanisms of these scaffolding approaches:

Research Question 1: What is the effect of scaffolded academic reading comprehension skill instruction using Gibbon's (2017) model on the reading achievement of the university students with regard to the motivational selves, learning resilience and self-regulation?

Research Question 2: What is the effect of scaffolded academic reading comprehension skill instruction using van del Pol et al.'s (2010) model on the reading achievement of the university students with regard to the motivational selves, learning resilience and self-regulation?

Research Question 3: What is the effect of Coyle's (2015) scaffolding approach when discussing reading achievement on the reading achievement of the university students with regard to the motivational selves, learning resilience and self-regulation?

Research Question 4: What mediation strategies teachers used in the scaffolded instructions in these three instructional contexts?

Method

Participants

The research encompassed 120 EFL learners, comprising both males and females, enrolled at the Islamic Azad University of Mazandaran, specifically in the branches located in Qaemshahr and Sari. These participants were undergraduate students pursuing their Bachelor's degrees and were selected through a convenience sampling method. The researchers included students studying in diverse academic disciplines such as humanities, social sciences, and engineering. The participants were between 19 and 25 years old, and all were native Persian speakers with comparable English language backgrounds, having completed at least one semester of general English courses as part of their curriculum. None of the participants had previously received formal instruction in scaffolded reading strategies. To uphold ethical standards, participants were thoroughly informed about the study's objectives and procedures, and explicit informed consent was secured before their involvement. Participation was entirely voluntary, with clear guarantees of confidentiality and the option to withdraw at any time.

Table 1

Demographic Characteristics of the Participants (N = 120)

Variable	Category	n	%
Gender	Male	58	48.3
	Female	62	51.7
Age Group	19–20 years	38	31.7
	21–23 years	52	43.3
	24–25 years	30	25.0
Educational Level	BA	120	100
	MA	0	0
	PhD	0	0

Materials and Instruments

Oxford Placement Test (OPT)

The Oxford Placement Test (OPT) has two sections, in which the grammar and vocabulary component typically comprises approximately 30 to 40 multiple-choice items completed within an allocated time of about 30 minutes, while the optional listening component includes roughly 15 to 20 items and requires an additional 15 minutes. The OPT test is a standardized assessment designed to evaluate English language proficiency, commonly used to place learners into appropriate language courses based on their skills in grammar, vocabulary, and listening. The test is adaptive, meaning the difficulty of questions adjusts based on the test-taker's responses. It has demonstrated strong psychometric properties, with high reliability coefficients (Cronbach's alpha > 0.80) and solid construct validity, effectively assessing general English proficiency across core language areas. In the present study, only the grammar and vocabulary sections of the OPT were administered to assess the homogeneity of participants prior to the main intervention. The listening section was omitted. The decision not to use the full OPT, which includes a reading section, was based on practical and methodological considerations. Specifically, the primary focus of this study was on reading achievement and emotion-related variables rather than comprehensive language proficiency, and the grammar and vocabulary subtests were deemed sufficient to control for participants' baseline language knowledge relevant to the study’s aims. The present study, only the grammar and vocabulary section were used, and the listening component was excluded. This decision was made to reduce overall testing time and to minimize participant fatigue, thereby lowering the time burden on respondents while still ensuring an accurate estimation of general English proficiency suitable for research purposes. Although the full OPT has established validity and reliability, prior research has also reported acceptable psychometric properties for the individual subtests when used independently. The grammar and vocabulary sections, in particular, have shown high internal consistency and strong correlations with overall English proficiency. Thus, their use as standalone measures remains methodologically sound. Nevertheless, the limitations of using a truncated version of the OPT are acknowledged, and results are interpreted with caution to account for any potential reduction in the comprehensiveness of proficiency assessment.

L2 Motivational Selves Questionnaire

The questionnaire, developed by Moskovsky et al. (2016), is designed to assess learners' motivation in language classrooms. It typically includes 15 items across three key components. All items are rated on a 5-point Likert scale. Earlier research has shown strong internal consistency, with Cronbach’s alpha values between 0.80 and 0.90, and test-retest reliability coefficients exceeding 0.75, reflecting high stability over time. Construct validity has also been supported by significant correlations with related motivation constructs (e.g., L2 Learning Motivation Scale, Self-Regulated Learning Strategies Scale), confirming that it effectively captures core elements of L2 motivation as proposed in Dörnyei and Csizér’s theoretical framework (2016). Considering the participants’ level of English proficiency, the original English version of the questionnaire was translated into Persian to ensure full comprehension and accuracy of responses. The translation process followed established procedures, including back-translation by independent bilingual experts. To ensure content and construct validity, a panel of experts in applied linguistics reviewed both versions. In addition, a pilot study was conducted with a representative sample of participants to assess item clarity and cultural appropriateness. Based on the pilot data, Cronbach’s alpha coefficients for the Persian version ranged from 0.82 to 0.88 across the three subscales, confirming strong internal consistency. The translated version maintained the original factor structure, as confirmed by confirmatory factor analysis (CFA), supporting the validity of the Persian version. The same translation and validation procedure was employed for the other two instruments used in this study to ensure consistency in linguistic and conceptual equivalence. This approach minimized measurement error due to language misunderstanding and strengthened the overall reliability and validity of the data collected (See Appendix A).

Resilience Questionnaire

The Questionnaire (RQ-26) is designed to measure an individual's capacity to cope with stress, adapt to challenges, and recover from adversity (Gartland et al., 2011). It typically assesses psychological resilience across multiple dimensions, such as emotional regulation, self-efficacy, optimism, and social support. The questionnaire usually follows a 5-point Likert scale. It consists of 10 items and five key dimensions. Emotional Regulation (2 items), Self-Efficacy (2 items), Optimism (2 items), Social Support (2 items), and Perseverance (2 items). Cronbach's alpha values range from 0.85 to 0.92, indicating high reliability. Studies report correlations above 0.80 when retested after a few weeks, showing stable results over time. Developed based on established resilience theories (Connor & Davidson, 2003; Masten, 2001), ensuring it covers key aspects of resilience (See Appendix B).

Self-regulation Questionnaire

The questionnaire is a commonly utilized tool developed to evaluate a person's capacity to manage their thoughts, emotions, and actions in order to attain long-term objectives (Carey et al., 2004). It measures self-regulation as a multidimensional construct involving goal setting, impulse control, and adaptability. The questionnaire consists of 20 items and five key dimensions, with responses typically rated on a 5-point Likert scale: Goal Setting (4 items), Self-Monitoring (4 items), Self-Control (4 items), Decision-Making (4 items), and Adaptability (4 items). Cronbach's alpha values range from 0.85 to 0.91, indicating strong reliability. Reported correlations above 0.80 over a period of weeks suggest high stability over time. Developed based on established theories of self-regulation (Zimmerman, 2000; Baumeister & Vohs, 2007), ensuring it covers key aspects of self-regulation (See Appendix C).

Reading Comprehension Test

The test is designed to assess an learners' ability to understand, interpret, and analyze written texts. The test consists of 30 multiple-choice questions based on different passages, covering a range of cognitive skills such as literal comprehension, inferential reasoning, and critical analysis. The test includes five reading passages, each followed by 6 multiple-choice questions. The passages vary in complexity and genre, including narrative, expository, argumentative, and informational texts. Each correct answer is awarded one point (Total: 30 points): 25–30: Advanced comprehension skills, 18–24: Proficient comprehension, 10–17: Basic comprehension, needs improvement, and 0–9: Below basic, significant improvement needed. Cronbach's alpha: 0.85 – 0.92, indicating high reliability. Correlation: 0.80+, demonstrating consistent results over time. Developed based on reading comprehension models (Kintsch, 1998; Snow, 2002) to cover key skills.

Procedure

This study adopted a pretest–posttest mixed-methods design to examine the effects of scaffolded academic reading instruction on EFL university students' reading achievement, motivation, resilience, and self-regulation. Both quantitative and qualitative data were collected to provide a comprehensive account of learning outcomes and learners' experiences. A total of 120 male and female EFL students were recruited from the Qaemshahr and Sari branches of Islamic Azad University, Mazandaran. Following participant recruitment, learners were randomly assigned to three equal experimental groups, each receiving instruction grounded in a distinct scaffolding framework.

Before the instructional phase, all participants took the Oxford Placement Test (OPT) to confirm comparability in general English proficiency across groups. In addition, baseline data were collected through four standardized instruments: a Reading Comprehension Test (30 items), the L2 Motivational Selves Questionnaire (15 items), a Resilience Questionnaire (26 items), and a Self-Regulation Questionnaire (32 items). The intervention phase lasted between eight and ten weeks, with two instructional sessions of 90 minutes per week. All groups worked with the same academic reading materials, but the instructional procedures differed based on the assigned scaffolding model. To maintain procedural consistency, instructors received prior training specific to the model they implemented, and instructional fidelity was monitored through classroom observations and systematic field notes.

Implementation of the Gibbons (2017) Scaffolding Model

In the first experimental group, reading instruction was organized according to Gibbons' (2017) scaffolding framework, which emphasizes gradual movement from supported participation to independent performance. Instruction began with high levels of teacher guidance, including explicit modeling of reading strategies, joint text construction, and guided discussions to activate background knowledge. Teachers made frequent use of think-aloud projects, visual organizers, and structured questioning to support learners' comprehension of academic texts. As instruction progressed, scaffolding was intentionally reduced, allowing students to assume greater responsibility for strategy use, interpretation, and critical engagement with texts. This gradual release aimed to foster autonomy while maintaining cognitive challenge.

Implementation of the van de Pol et al. (2010) Scaffolding Model

The second group received instruction based on the contingent scaffolding model proposed by van de Pol et al. (2010). In this approach, instructional support was continuously adjusted in response to learners' immediate needs. Teachers closely monitored students' performance during reading tasks and provided prompts, feedback, or clarification only when difficulties emerged. Scaffolding was therefore dynamic rather than predetermined, with support withdrawn or intensified depending on the students' demonstrated understanding. This model prioritized diagnostic interaction, emphasizing sensitivity to learner responses and moment-by-moment instructional decision-making during academic reading activities.

Implementation of the Coyle (2015) Scaffolding Model

Instruction in the third group followed Coyle's (2015) framework, which integrates scaffolding within a content-and-language-integrated learning orientation. Reading lessons were structured around meaningful academic themes, and scaffolding focused on supporting both content understanding and language development. Teachers employed pre-reading tasks to build conceptual and lexical readiness, during-reading activities to guide comprehension and analysis, and post-reading tasks to encourage reflection and application. Scaffolding strategies included structured peer interaction, task sequencing from simple to complex, and explicit attention to academic language forms relevant to the texts. This approach aimed to situate reading within a broader learning context while sustaining learner engagement.

Following the completion of the instructional phase, all participants completed the same set of questionnaires and the reading comprehension test as post-tests. To complement the quantitative findings, qualitative data were collected through semi-structured interviews with a purposive sample of five to seven students from each group. The interviews explored learners' experiences with the scaffolded instruction, perceived benefits and challenges, and comparisons with prior reading instruction. The interview protocol was aligned with the constructs measured in the questionnaires, allowing for triangulation between quantitative outcomes and students' reported perceptions.

Data Analysis

Quantitative data were examined through both descriptive statistics and inferential analysis methods, including MANOVA to examine the effects of the instructional models on multiple dependent variables—reading achievement, motivation, resilience, and self-regulation. Effect sizes were also calculated. Qualitative data were analyzed thematically, with a focus on identifying patterns of mediation strategies and learners' perceptions. Thematic coding was applied to both interview transcripts and classroom interaction episodes. To ensure validity and reliability, strategies such as instrument standardization, observer triangulation, member checking, and pilot testing of the interview protocol were employed. Ethical considerations were addressed by obtaining informed consent from participants, assuring anonymity and the right to withdraw, and securing institutional ethical approval prior to data collection.

Results

Before conducting the MANOVA for RQ1, several assumptions were tested to ensure the validity of the multivariate analysis. This included normality, homogeneity of variance-covariance matrices, multicollinearity, and absence of outliers. To assess univariate normality, Shapiro–Wilk tests were conducted for each dependent variable within each group (Gibbon's and Control). The results are summarized below:

Before conducting the MANOVA for RQ1, a series of assumption checks were carried out to ensure the validity of the multivariate analysis. The results confirm that all statistical assumptions for conducting a MANOVA were sufficiently met. The SPSS MANOVA output for RQ1 is presented in the following tables.

Table 2

Descriptive Statistics for the Gibbon's Groups

Group		Reading Achievement	Motivational Selves	Resilience	Self-Regulation
Gibbon's	0	25.10 (SD = 2.15)	61.30 (SD = 5.40)	80.40 (SD = 4.90)	95.10 (SD = 6.20)
Control	0	21.20 (SD = 2.50)	56.80 (SD = 4.95)	75.20 (SD = 5.30)	88.50 (SD = 5.80)

Descriptive statistics for the Gibbon's and Control groups are presented in Table 2. The Gibbon's group (N = 30) demonstrated higher mean scores than the Control group (N = 30) on reading achievement (M = 25.10, SD = 2.15 vs. M = 21.20, SD = 2.50), motivational selves (M = 61.30, SD = 5.40 vs. M = 56.80, SD = 4.95), resilience (M = 80.40, SD = 4.90 vs. M = 75.20, SD = 5.30), and self-regulation (M = 95.10, SD = 6.20 vs. M = 88.50, SD = 5.80). These descriptive results suggest that participants in the Gibbon's group outperform the Control group across cognitive, motivational, and self-regulatory domains.

Table 3

Multivariate Tests for the Gibbon's Groups

Effect	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Group	Pillai's Trace = 0.637	11.211	4	55	.000	.637

The results of Table 3 report that a MANOVA was conducted to scrutinize the effect of Group on the combined dependent variables. The multivariate test using Pillai's Trace indicated a significant effect of Group, Pillai's Trace = 0.637, $F(4, 55) = 11.21$, $p < .001$, with a large effect size (Partial $\eta^2 = .637$). These results suggest that the groups differ significantly when considering the dependent variables together.

Table 4

Tests of Between-Subjects Effects for the Gibbon's Groups

Dependent Variable	Source	Type III SS	Mean Square	F	Sig.	Partial Eta Squared
Reading Achievement	Group	172.90	172.90	34.58	.000	.376
Motivational Selves	Group	288.80	288.80	10.45	.002	.155
Resilience	Group	484.50	484.50	19.31	.000	.248
Self-Regulation	Group	799.20	799.20	22.17	.000	.281

As shown in Table 4, the tests revealed significant differences between groups on all dependent variables. Group had a significant effect on reading achievement, $F(1, 58) = 34.58$, $p < .001$, Partial $\eta^2 = .376$; motivational selves, $F(1, 58) = 10.45$, $p = .002$, Partial $\eta^2 = .155$; resilience, $F(1, 58) = 19.31$, $p < .001$, Partial $\eta^2 = .248$; and self-regulation, $F(1, 58) = 22.17$, $p < .001$, Partial $\eta^2 = .281$. These results indicate that group membership accounts for a substantial proportion of variance across these psychological and academic outcomes.

Table 5

Multivariate Tests for the Gibbon's Groups

Variable	Gibbon's Mean	Control Mean	Sig.
Reading Achievement	25.10	21.20	.000
Motivational Selves	61.30	56.80	.002
Resilience	80.40	75.20	.000
Self-Regulation	95.10	88.50	.000

As presented in Table 5, the estimated marginal means indicate that the Gibbon's group scored significantly higher than the Control group across all measured variables. Specifically, reading achievement (M = 25.10 vs. 21.20), motivational selves (M = 61.30 vs. 56.80), resilience (M = 80.40 vs. 75.20), and self-regulation (M = 95.10 vs. 88.50) all showed statistically significant differences, with p-values of less than .01. These findings suggest that the Gibbon's instructional approach is associated with enhanced academic and psychological outcomes compared to the Control condition. The MANOVA results indicate a statistically significant multivariate effect of Gibbon's scaffolded instruction model on the combined dependent variables. Follow-up univariate tests show significant differences across all four outcomes — reading achievement, motivational selves, resilience, and self-regulation — favoring the Gibbon's model group.

Before conducting the MANOVA for RQ2, a series of assumption checks were carried out to ensure the validity of the multivariate analysis. The results confirm that all statistical assumptions for conducting a MANOVA were sufficiently met. The SPSS MANOVA outputs for RQ2 are presented in the following tables.

Table 6

Descriptive Statistics for the van de Pol Groups

Group		Reading Achievement	Motivational Selves	Resilience	Self-Regulation
van de Pol	30	24.30 (SD = 2.30)	60.10 (SD = 5.00)	78.60 (SD = 4.80)	93.70 (SD = 6.00)
Control	30	21.20 (SD = 2.50)	56.80 (SD = 4.95)	75.20 (SD = 5.30)	88.50 (SD = 5.80)

Table 6 presents descriptive statistics for the van de Pol and Control groups. The van de Pol group demonstrated higher mean scores across all measured variables compared to the Control group. Specifically, reading achievement (*M* = 24.30 vs. 21.20), motivational selves (*M* = 60.10 vs. 56.80), resilience (*M* = 78.60 vs. 75.20), and self-regulation (*M* = 93.70 vs. 88.50) were all elevated in the van de Pol group, indicating potential benefits of the van de Pol instructional approach on both academic and psychological outcomes.

Table 7

Multivariate Tests for the van de Pol Groups

Effect	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Group	Pillai's Trace = 0.559	8.729	4	55	.000	.559

Table 7 illustrates that the multivariate test found a statistically significant effect of Group on the combined dependent variables, with Pillai's Trace = 0.559, $F(4, 55) = 8.73$, $p < .001$, and a partial η^2 of .559. This suggests that group membership explains a considerable amount of the variance across all dependent variables taken together.

Table 8

Tests of Between-Subjects Effects for the van de Pol Groups

Dependent Variable	Source	Type III SS	f	Mean Square	F	Sig.	Partial Eta Squared
Reading Achievement	Group	136.80	1	136.80	24.50	.000	.308
Motivational Selves	Group	216.45	1	216.45	7.91	.007	.126
Resilience	Group	312.30	1	312.30	12.77	.001	.188
Self-Regulation	Group	599.80	1	599.80	17.30	.000	.239

Table 8 displays the results of tests of between-subjects effects for the influence of Group on reading achievement, motivational selves, resilience, and self-regulation. Significant group differences were found across all dependent variables: reading achievement, $F(1, 58) = 24.50$, $p < .001$, partial $\eta^2 = .308$; motivational selves, $F(1, 58) = 7.91$, $p = .007$, partial $\eta^2 = .126$; resilience, $F(1, 58) = 12.77$, $p = .001$, partial $\eta^2 = .188$; and self-regulation, $F(1, 58) = 17.30$, $p < .001$, partial $\eta^2 = .239$. These results indicate that group membership significantly influences all measured outcomes, with effect sizes ranging from moderate to large.

Table 9

Estimated Marginal Means for the van de Pol Groups

Variable	van de Pol Mean	Control Mean	Sig.
Reading Achievement	24.30	21.20	.000
Motivational Selves	60.10	56.80	.007
Resilience	78.60	75.20	.001
Self-Regulation	93.70	88.50	.000

As shown in Table 9, estimated marginal means indicate statistically significant differences between the van de Pol and control groups across all measured outcomes. Specifically, the van de Pol group demonstrated higher scores in reading achievement ($M = 24.30$ vs. 21.20, $p < .001$), motivational selves (M = 60.10 vs. 56.80, $p = .007$), resilience ($M = 78.60$ vs. 75.20, $p = .001$), and self-regulation ($M = 93.70$ vs. 88.50, $p < .001$). These findings support the effectiveness of the van de Pol scaffolding model in enhancing both academic and psychological learner characteristics. There is a statistically significant multivariate effect of the van de Pol scaffolded instruction model on the combined dependent variables. Univariate follow-up tests indicate significant improvements in reading achievement, motivational selves, resilience, and self-regulation in the van de Pol group compared to the control group.

Before conducting the MANOVA for RQ3, a series of assumption checks were carried out to ensure the validity of the multivariate analysis. The results confirm that all statistical assumptions for conducting a MANOVA were sufficiently met. The SPSS MANOVA output for RQ3 is presented in the following tables.

Table 10

Descriptive Statistics for Coyle's Groups

Group	**Reading Achievement**	**Motivational Selves**	**Resilience**	**Self-Regulation**
Coyle 0	25.00 (SD = 2.00)	61.50 (SD = 4.80)	80.00 (SD = 4.50)	95.20 (SD = 5.50)
Control 0	21.10 (SD = 2.40)	56.70 (SD = 5.10)	75.00 (SD = 5.10)	88.40 (SD = 5.60)

As presented in Table 10, the descriptive statistics revealed that the group taught through Coyle's scaffolded instruction model consistently outperformed the control group across all measured domains. Specifically, the Coyle group showed higher mean scores in reading achievement (M = 25.00, SD = 2.00), motivational selves (M = 61.50, SD = 4.80), resilience (M = 80.00, SD = 4.50), and self-regulation (M = 95.20, SD = 5.50) compared to the control group (M = 21.10, 56.70, 75.00, and 88.40 respectively). These results suggest a notable positive impact of Coyle's instructional approach on both academic and psychological learner characteristics.

Table 11

Multivariate Tests for Coyle's Groups

Effect	**Value (Pillai's Trace)**	**F**	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Group	0.601	10.370	4	55	.000	.601

As shown in Table 11, the multivariate test using Pillai's Trace discovered a statistically significant effect of group on the combined dependent variables—reading achievement, motivational selves, resilience, and self-regulation, Pillai's Trace = 0.601, $F(4, 55) = 10.37$, $p < .001$, partial $\eta^2 = .601$. This indicates that group membership (Coyle vs. Control) accounted for a substantial proportion of variance in the outcome measures.

Table 12

Multivariate Tests for Coyle's Groups

Dependent Variable	Source	Type III SS	Mean Square	F	Sig.	Partial Eta Squared
Reading Achievement	Group	170.10	170.10	30.50	.000	.357
Motivational Selves	Group	240.25	240.25	9.80	.003	.151
Resilience	Group	345.60	345.60	14.20	.000	.205
Self-Regulation	Group	648.00	648.00	19.60	.000	.263

Table 12 indicates that there were meaningful statistical differences between the Coyle and Control groups across all four dependent measures. The group variable significantly influenced outcomes in reading achievement (F(1, df) = 30.50, $p < .001$, partial $\eta^2 = .357$), motivational selves (F = 9.80, $p = .003$, partial $\eta^2 = .151$), resilience (F = 14.20, $p < .001$, partial $\eta^2 = .205$), and self-regulation (F = 19.60, $p < .001$, partial $\eta^2 = .263$). These findings demonstrate that the Coyle group performed better than the Control group in each domain, with effect sizes ranging from moderate to substantial.

Table 13

Multivariate Tests for Coyle's Groups

Variable	Coyle Mean	Control Mean	Sig.
Reading Achievement	25.00	21.10	.000
Motivational Selves	61.50	56.70	.003
Resilience	80.00	75.00	.000
Self-Regulation	95.20	88.40	.000

As shown in Table 13, the estimated marginal means revealed statistically significant group differences favoring the Coyle group over the Control group across all outcome measures. Specifically, the Coyle group scored significantly higher on reading achievement ($M = 25.00$ vs. 21.10, $p < .001$), motivational selves ($M = 61.50$ vs. 56.70, $p = .003$), resilience ($M = 80.00$ vs. 75.00, $p < .001$), and self-regulation ($M = 95.20$ vs. 88.40, $p < .001$). These results suggest that the Coyle et al. scaffolded instruction model positively impacted learners' academic and motivational-emotional outcomes. The results show a statistically significant multivariate effect of Coyle et al.'s (2015) scaffolded instruction model on all four dependent variables. Univariate analyses revealed significantly higher reading achievement, motivational selves, resilience, and self-regulation scores for the experimental group using the Coyle model compared to the control group.

Below is the complete answer using MAXQDA coding and thematic analysis for the research question. This includes thematic codes, examples, and extracts from the interviews conducted with students who implemented scaffolded academic reading instruction based on the three models Gibbons, van de Pol et al., and Coyle et al.

Table 14

Main Themes (Mediation Strategies Identified)

Code (Theme)	Description	Frequency
Modeling Cognitive Processes	Teachers demonstrated strategies for reading comprehension (thinking aloud)	21
Prompting and Cueing	Use of verbal prompts or hints to guide student responses	26
Feedback and Recasting	Corrective and formative feedback that scaffolds the learning process	18
Gradual Release of Responsibility	Shift from teacher-led to student-led comprehension	15
Clarification Requests	Encouraging students to explain or clarify their understanding	12
Use of Graphic Organizers	Scaffolding through visual mapping tools (e.g., KWL charts, mind maps)	11
Building Background Knowledge	Activating or supplementing prior knowledge	14
Language Support Scaffolds	Lexical or grammatical support (sentence starters, vocabulary previews)	19

As shown in Table 14, qualitative analysis of classroom interactions revealed eight primary mediation strategies employed by teachers. Prompting and cueing emerged as the most frequently identified theme ($f = 26$), followed by modeling cognitive processes ($f = 21$) and language support scaffolds ($f = 19$). These strategies reflect a scaffolding-oriented pedagogy aimed at supporting students' reading comprehension and autonomy. Other notable strategies included feedback and recasting ($f = 18$), the gradual release of responsibility ($f = 15$), and building background knowledge ($f = 14$). Collectively, these findings underscore the multifaceted nature of instructional scaffolding in promoting literacy development and learner engagement.

Theme 1: Modeling Cognitive Processes

"When we started the session, I would read a paragraph and say aloud what I was thinking. For example, I'd say, 'This sentence has a cause-effect relationship because of the phrase due to.' This helped them see how to analyze text." (Student C, Gibbons-based instruction)

Coding: Modeling → Think-Aloud Strategy
Interpretation: Modeling as a cognitive apprenticeship was a dominant scaffold in all three models but most systematically used in Gibbons' model.

Theme 2: Prompting and Cueing

"I didn't tell them the answer. Instead, I'd ask, 'Does the next sentence give more details or contrast the idea?' This made them read more attentively." (Student B, van de Pol model)

Coding: Prompting → Metacognitive Cue
Interpretation: Prompts activated metacognitive thinking; van de Pol model emphasized contingent support with prompts adjusted to student readiness.

Theme 3: Feedback and Recasting

"When a student answered incorrectly, I'd rephrase it slightly and ask, 'Is this what you meant?' This way they felt safe and guided." (Student A, Coyle model)

Coding: Feedback → Recasting
Interpretation: Teachers used scaffolding via feedback loops to support language accuracy and conceptual understanding.

Theme 4: Gradual Release of Responsibility

"At first, I guided them through the comprehension questions. Later, I stepped back and just monitored their group discussions." (Student C, Gibbons model)

Coding: Scaffolding Structure → Gradual Release
Interpretation: Gibbons and Coyle models emphasized moving from guided instruction to independent student work.

Theme 5: Clarification Requests

"When they misunderstood a text, I didn't explain right away. I asked, 'Can you say what this paragraph means in your own words?'" (Student B, van de Pol model)

Coding: Interaction → Clarification Request
Interpretation: Encouraged deeper processing and peer explanations, aligned with responsive mediation.

Theme 6: Use of Graphic Organizers

"I gave them a cause-effect chart before reading. This helped them trace the logic while reading." (Student A, Coyle model)

Coding: Visual Scaffold → Organizer Use
Interpretation: Visual tools were instrumental in Coyle's CLIL model, aiding comprehension of complex texts.

Theme 7: Building Background Knowledge

"Before the reading on environmental issues, I showed a short video and we discussed pollution in our city." (Student C, Gibbons model)

Coding: Pre-Reading → Schema Activation
Interpretation: Essential for contextualizing academic texts; strongly emphasized in all models.

Theme 8: Language Support Scaffolds

"Some words were tough, so I gave them a mini-glossary and sentence starters like 'The author argues that…'." (Student A, Coyle model)

Coding: Language Scaffold → Sentence Frame
Interpretation: Support for both linguistic and academic discourse features was most developed in Coyle's model.

Table 15

Cross-Model Summary of Mediation Strategies

Mediation Strategy	Gibbons	van de Pol	Coyle
Modeling	✓✓✓	✓	✓
Prompting and Cueing	✓✓	✓✓✓	✓✓
Feedback and Recasting	✓	✓✓	✓✓✓
Gradual Release of Responsibility	✓✓✓	✓	✓✓
Clarification Requests	✓	✓✓✓	✓
Graphic Organizers	✓	✓	✓✓✓
Background Knowledge Activation	✓✓✓	✓	✓
Language Support	✓✓	✓	✓✓✓

✓ = Present | ✓✓✓ = Strongly present

As illustrated in Table 15, all three instructional models—Gibbons, van de Pol, and Coyle—employ a range of mediation strategies, albeit with varying emphases. Gibbons' model strongly integrates modeling cognitive processes, gradual release of responsibility, and background knowledge activation, aligning with its scaffolding philosophy. In contrast, van de Pol's model exhibits a stronger orientation toward prompting and cueing, clarification requests, and feedback, emphasizing responsive interaction. Coyle's model is distinguished by a robust use of graphic organizers, feedback and recasting, and language support scaffolds, suggesting an emphasis on visual and linguistic mediation. These patterns underscore the models' differential approaches to scaffolding, each contributing uniquely to enhancing student comprehension and engagement. Teachers utilized a range of mediation strategies across the three instructional models. While all strategies were present to some extent in each context, modeling and background activation were more pronounced in Gibbons' model; prompting, feedback, and clarification dominated van de Pol's responsive scaffolding approach; and visual tools and language scaffolds were most elaborately employed under Coyle's CLIL framework. These differences reflect the underlying pedagogical emphasis of each model.

Discussion

The findings from the MANOVA analyses conducted for Research Questions 1, 2, and 3 provide compelling and converging evidence for the effectiveness of scaffolded academic reading comprehension instruction, grounded in the theoretical frameworks proposed by Gibbons (2017), van de Pol et al. (2010), and Coyle (2015). Each of these models, although distinct in their pedagogical underpinnings, shares a common commitment to providing structured, responsive, and gradual support that facilitates learners' transition from assisted to independent performance. For RQ1, which examined the impact of Gibbons's model, descriptive statistics revealed that students in the experimental group ($N = 30$) demonstrated consistently higher outcomes across all measured domains compared to the control group ($N = 30$). The multivariate analysis confirmed a statistically significant overall effect, with a large effect size (partial $\eta^2 = .637$), indicating that the intervention produced substantial gains across cognitive and non-cognitive variables. Follow-up univariate ANOVAs reinforced this conclusion, revealing large effect sizes for reading achievement ($\eta^2 = .376$) and self-regulation ($\eta^2 = .281$), and moderately large effects for resilience ($\eta^2 = .248$) and motivational selves ($\eta^2 = .155$). These findings are consistent with Gibbons's sociocultural approach, which emphasizes scaffolding language and content simultaneously through contextualized, dialogic teaching practices. The model appears to promote not only academic reading success but also learners' emotional and motivational capacity to persist and thrive in challenging academic environments.

Likewise, the results for RQ2, which assessed the instructional model proposed by van de Pol et al. (2010), further validate the pedagogical value of scaffolded instruction in EFL contexts. Univariate tests confirmed statistically significant gains across all four outcome variables, with large effect sizes for reading achievement ($\eta^2 = .308$) and self-regulation ($\eta^2 = .239$), and moderately strong effects for resilience ($\eta^2 = .188$) and motivational selves ($\eta^2 = .126$). Van de Pol's model is rooted in a dynamic scaffolding process, where teachers continually adjust their support in response to learners' evolving needs. These results suggest that such a responsive teaching approach not only boosts reading comprehension but also enhances learners' emotional resilience and capacity for autonomous learning. Although the overall effect sizes were slightly lower than those of the Gibbons model, the findings still underscore the significant pedagogical benefits of incorporating adaptive scaffolding mechanisms into classroom instruction.

In the case of RQ3, the analysis centered on Coyle's (2015) model, grounded in the 4Cs framework, which integrates content, communication, cognition, and culture in a holistic instructional design. The results similarly demonstrated that students exposed to Coyle's scaffolded instruction outperformed those in the control group across all dependent variables. Follow-up univariate tests showed statistically significant and practically meaningful differences on all four outcomes. These results provide empirical support for the CLIL approach, suggesting that an integrated focus on academic content and language development can foster deeper cognitive engagement and emotional investment, which are critical for sustained academic success and learner well-being. In addition to enhancing reading achievement, Coyle's model appears to nurture learners' identity, resilience, and self-directed learning—dimensions that are often neglected in traditional EFL instruction.

Taken together, these findings underscore the transformative potential of scaffolded academic reading instruction, particularly when informed by well-articulated theoretical models. While all three interventions yielded significant improvements in students' cognitive and affective outcomes, the Gibbons model emerged as the most effective overall, followed closely by Coyle's and then van de Pol's models. The consistency of results across different scaffolding frameworks reinforces the broader pedagogical principle that structured, adaptive, and context-sensitive support is critical in helping EFL learners meet the complex demands of academic literacy. These results carry significant weight for shaping curricula, improving teacher development programs, and informing educational policy—especially in contexts where students encounter both language-related and cognitive hurdles. Implementing scaffolded instructional strategies in a structured way can foster growth not just in reading comprehension, but also in students' emotional readiness, drive to learn, and ability to thrive as lifelong learners.

The results of this research offer strong empirical evidence that scaffolded instructional approaches effectively bolster both cognitive performance and affective-motivational aspects for EFL learners. These outcomes are consistent with a wide array of previous studies, reinforcing the pedagogical value of structured support in language education. Specifically, the significant improvements observed in reading achievement, motivational selves, resilience, and self-regulation across all three scaffolded instruction models validate the claims of sociocultural theorists such as Vygotsky, who emphasized the critical role of guided interaction and mediated learning within the learner's zone of proximal development (ZPD). Gibbons' (2017) framework, which centers on the integration of academic language and disciplinary content through scaffolded classroom discourse, produced the highest effect sizes in the present study—a finding that echoes earlier research by Acosta-Gonzaga and Ramirez-Arellano (2022), who reported that well-structured scaffolding strategies help English language learners (ELLs) access cognitively demanding tasks and build academic autonomy.

Furthermore, these results are consistent with van de Pol et al.'s (2010) review, which highlighted the importance of adaptive scaffolding practices that respond to students' evolving needs in real time. Although the van de Pol model in this study demonstrated slightly lower outcomes compared to the Gibbons and Coyle models, it still resulted in significant gains across all dependent variables, affirming the importance of contingent support as emphasized in classroom-based research by Mohammadi et al. (2020). Their work demonstrated that teacher responsiveness and scaffolding adjustments lead to increased student participation and task persistence. Similarly, research by Uçak and Kartal (2023) emphasized that scaffolding facilitates deeper conceptual understanding when it targets both content and process, which is evident in the performance of learners across the three models in this study. The effectiveness of teachers' support behaviour hinges on building their identity perceptions as a part of teachers' supportive professional cognition (Rostami et al., 2021).

The positive outcomes associated with Coyle's (2015) Content and Language Integrated Learning (CLIL) framework also align with previous CLIL-focused research that links integrated instruction to heightened learner engagement, improved academic language proficiency, and greater content retention. Studies by Ahmed Abdel-Al Ibrahim et al. (2023), as well as Teng et al. (2024), found that CLIL environments foster increased motivation, deeper cognitive processing, and emotional engagement—findings mirrored in this study's observed gains in motivational selves and resilience under the CLIL-based scaffolding condition. Dalton-Puffer (2011) similarly noted that CLIL supports higher-order thinking and learner autonomy, especially when scaffolding is explicitly embedded into content instruction. Moreover, the inclusion of affective outcomes such as resilience and self-regulation in the present study expands the scope of typical CLIL investigations, which often focus primarily on language and content outcomes (Wang et al., 2024).

Additional studies further corroborate these findings. For example, Uçak and Kartal (2023) emphasized that scaffolding that incorporates dialogic instruction and academic language development promotes identity formation and learner confidence—benefits that were most strongly observed in the Gibbons model group. Similarly, studies by Wang et al. (2024) demonstrated that structured dialogic teaching improves both cognitive and socio-emotional outcomes, supporting the link between language use and emotional resilience seen in this study. In terms of motivational selves and self-regulation, the findings are in line with Michalsky (2021), which posits that learner engagement increases when instructional practices activate learners' ideal L2 selves—a mechanism likely supported through personalized, scaffolded instruction that challenges learners while providing structured support.

Importantly, the current study extends prior research by simultaneously examining multiple emotional and motivational constructs within a scaffolded instructional context. While earlier studies (Ahmed Abdel-Al Ibrahim et al., 2023; Alfares, 2025) have typically focused on either academic achievement or isolated affective outcomes, this study offers a more holistic view of learner development, emphasizing the interplay between cognitive performance and psychological resilience. This aligns with contemporary educational psychology perspectives (Michalsky, 2021), which advocate for emotionally supportive instructional environments that balance cognitive challenge with affective scaffolding. Taken together, the present findings not only validate the pedagogical principles underlying scaffolded instruction but also highlight its multifaceted benefits in EFL contexts. The alignment with and expansion upon existing research underscore the need for teachers to implement theoretically informed, flexible, and emotionally responsive scaffolding practices that nurture both academic competence and emotional resilience in language learners. As educational systems increasingly emphasize the development of 21st-century skills—including self-regulation, persistence, and cross-cultural communication—such evidence-based instructional strategies become ever more essential.

Conclusion

This research presents persuasive support for the positive impact of scaffolded teaching methods in improving EFL learners' academic achievement as well as their emotional and motivational engagement. By comparing three prominent models—Gibbons' language scaffolding framework, van de Pol et al.'s contingent scaffolding approach, and Coyle's CLIL-based model—the findings reveal that scaffolded instruction significantly improves reading achievement, motivational selves, resilience, and self-regulation. Among the models, Gibbons' approach demonstrated the highest overall impact, particularly in fostering emotional resilience and self-regulated learning strategies. The results affirm the theoretical foundations of sociocultural theory and contemporary educational psychology, underscoring the value of structured, responsive, and contextually relevant support in language learning. Importantly, this study expands the scope of scaffolded instruction research by highlighting its emotional and motivational benefits alongside cognitive gains, thereby emphasizing the holistic development of learners. As such, the integration of scaffolded pedagogies into EFL instruction offers a promising pathway for cultivating competent, motivated, and emotionally resilient language learners. These insights have practical implications for curriculum design, teacher training, and classroom practice in diverse educational settings.

This research underscores both the conceptual and applied importance of scaffolded instruction within EFL environments, drawing from sociocultural theory, self-determination theory, and the L2 motivational self-system. The findings demonstrate that structured support strategies foster improvements in cognitive domains—such as reading comprehension—and affective-motivational areas including resilience, self-regulation, and learner identity. These outcomes affirm Vygotsky's notion of the Zone of Proximal Development by illustrating how targeted guidance facilitates learner growth. The findings emphasize the importance of emotionally attuned, socially interactive scaffolding in shaping learners' motivation and academic self-concept. Practically, the study recommends structured strategies like modeling, guided practice, and emotional support to reduce anxiety and foster engagement. It calls for teacher training and curriculum design that incorporate both cognitive and emotional scaffolding to address the diverse needs of EFL learners.

While this study provides valuable insights into how scaffolded instruction influences both the academic performance and emotional development of EFL learners, it is not without limitations. The narrow and context-bound sample constrains the broader applicability of the results. Additionally, the brief duration of the intervention curtails the ability to evaluate long-term impacts, emphasizing the importance of longitudinal studies to examine enduring effects.

To advance understanding of scaffolded instruction in EFL contexts, future research should: (1) involve larger, more diverse samples to improve generalizability; (2) use longitudinal designs to assess long-term impacts on academic and emotional outcomes; (3) investigate the effects of scaffolding on a range of language skills beyond reading; and (4) explore how teacher-related factors influence the effectiveness of scaffolding, informing better training and professional development.

References

Abdelshaheed, B. S. M. (2023). Student versus expert outlines in reading comprehension: The effect of collaborative construction. *SAGE Open*, *13*(3), 21582440231191795. https://doi.org/10.1177/21582440231191795

Acosta-Gonzaga, E., & Ramirez-Arellano, A. (2022). Scaffolding matters? Investigating its role in motivation, engagement and learning achievements in higher education. *Sustainability*, *14*(20), 13419. https://doi.org/10.3390/su142013419

Ahmed Abdel-Al Ibrahim, K., Cuba Carbajal, N., Zuta, M. E. C., & Bayat, S. (2023). Collaborative learning, scaffolding-based instruction, and self-assessment: Impacts on intermediate EFL learners' reading comprehension, motivation, and anxiety. *Language Testing in Asia*, *13*(1), 16. https://doi.org/10.1186/s40468-023-00229-1

Alfares, N. (2025). Effectiveness of a scaffolded autonomous learning scheme in enhancing learners' motivation in EFL classes. *Learning and Motivation*, *89*, Article 102084. https://doi.org/10.1016/j.lmot.2024.102084

Allagui, B. (2024). A scaffolding intervention to improve self-efficacy in source-based argumentative writing. *Frontiers in Psychology*, *15*, Article 1454104. https://doi.org/10.3389/fpsyg.2024.1454104

Baranovskaya, T. (2015). Self-regulation skills: Several ways of helping students develop self-regulated learning. *Journal of Language and Education*, *1*(2), 56–64. https://doi.org/10.17323/2411-7390-2015-1-2-56-64

Carey, K. B., Neal, D. J., & Collins, S. E. (2004). A psychometric analysis of the self-regulation questionnaire. *Addictive behaviors*, *29*(2), 253–260. https://doi.org/10.1016/j.addbeh.2003.08.001

Coyle, D. (2015). Strengthening integrated learning: Towards a new era for pluriliteracies and intercultural learning. *Latin American Journal of Content & Language Integrated Learning*, *8*(2), 84–103. https://doi.org/10.5294/5915

Danli, L. (2017). Autonomy in scaffolding as learning in teacher-student negotiation of meaning in a university EFL classroom. *Chinese Journal of Applied Linguistics*, *40*(4), 410–430. https://doi.org/10.1515/cjal-2017-0024

Derakhshan, A., & Fathi, J. (2025). From boredom to buoyancy: Examining the impact of perceived teacher support on EFL learners' resilience and achievement through a serial mediation model. *Innovation in Language Learning and Teaching*, Advance Online Publication. https://doi.org/10.1080/17501229.2025.2487932.

Derakhshan, A., Setiawan, S., & Ghafouri, M. (2024). Modeling the interplay of Indonesian and Iranian EFL teachers' apprehension, resilience, organizational mattering, and psychological well-being. *Iranian Journal of Language Teaching Research*, *12*(1), 21–43. https://doi.org/10.30466/ijltr.2024.121416

Estrada-Araoz, E. G., Sayed, B. T., Niyazova, G. G., & Lami, D. (2023). Comparing the effects of computerized formative assessment vs. computerized dynamic assessment on developing EFL learners' reading motivation, reading self-concept, autonomy, and self-regulation. *Language Testing in Asia*, *13*(1), 39. https://doi.org/10.1186/s40468-023-00253-1

Gartland, D., Bond, L., Olsson, C. A., Buzwell, S., & Sawyer, S. M. (2011). Development of a multi-dimensional measure of resilience in adolescents: The adolescent resilience questionnaire. *BMC Medical Research Methodology*, *11*, 1–10. https://doi.org/10.1186/1471-2288-11-134

Grocott, L., McEntee, K., Coleman, K., & Manix, R. (2019). The becoming of a designer: An affective pedagogical approach to modelling and scaffolding risk-taking. *Art, Design & Communication in Higher Education*, *18*(1), 99–112. https://doi.org/10.1386/adch.18.1.99_1

Halterman, D. A. (2023). *Increasing reading comprehension: Continuing educator training to better teach the science of reading for English language learners.* PhD Dissertation, University of Massachusetts Lowell. https://www.proquest.com/openview/a429461cfe8b2aba7261d5c2323c324c/1?pq-origsite=gscholar&cbl=18750&diss=y

Härkki, T. (2020). Eye on collaborative creativity: Insights from multiple-person mobile gaze tracking in the context of collaborative design. the interdisciplinarity of the learning sciences, *14th International Conference of the Learning Sciences (ICLS) 2020: Proceedings Volume 5,* ISLS International Society of the Learning Sciences.

Hassen, S. K., Adugna, E. T., & Bogale, Y. N. (2024). EFL students' writing achievement via the execution of scaffolding strategies instruction. *System*, *125*, 103439. https://doi.org/10.1016/j.system.2024.103439

Hou, L. (2025). Unboxing the intersections between self-esteem and academic mindfulness with test emotions, psychological wellness and academic achievement in artificial intelligence-supported learning environments: Evidence from English as a foreign language learners. *British Educational Research Journal*, *n/a*(n/a). https://doi.org/https://doi.org/10.1002/berj.4146

Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning*, *64*(1), 160-212. https://doi.org/10.1111/lang.12034

Joyce-Gibbons, A. (2017). Observe, interact and act: Teachers' initiation of mini-plenaries to scaffold small-group collaboration. *Technology, Pedagogy and Education*, *26*(1), 51–68. https://doi.org/10.1080/1475939X.2016.1173089

Li, C.-T., Hou, H.-T., & Lee, L.-H. (2025). Design of a dual-hierarchy scaffolding board game-based learning activity for EFL reading comprehension. *Language Teaching Research*, *29*(6), 13621688221125903. https://doi.org/10.1177/13621688221125903

Mallahi, O. (2025). Exploring the relationship between Iranian EFL learners' L2 writing self (ideal and ought-to L2 writing selves), growth mindset, and their feedback-seeking behavior. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, *13*(2), 229-254. https://doi.org/10.22049/jalda.2025.30740.1823

Marzban, B., Adel, S. M. R., Eghtesadi, A. R., & Elyasi, M. (2021). Identity and its ramifications among advanced adult Iranian learners of English. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, *9*(2), 35-57. https://doi.org/10.22049/jalda.2021.27193.1288

Mellati, M., & Valizadeh, M. (2025). Exploring the emotional landscape: How graduate students' views on social networks influence their engagement in tech-enhanced learning environments. *Education and Information Technology,30, 16269*-16290. https://doi.org/10.1007/s10639-025-13613-6

Michalsky, T. (2021). When to scaffold motivational self-regulation strategies for high school students' science text comprehension. *Frontiers in Psychology*, *12*, Article 658027. https://doi.org/10.3389/fpsyg.2021.658027

Mohammadi, R. R., Saeidi, M., & Ahangari, S. (2020). Self-regulated learning instruction and the relationships among self-regulation, reading comprehension and reading problem solving: PLS-SEM approach. *Cogent Education*, *7*(1), Article 1746105. https://doi.org/10.1080/2331186X.2020.1746105

Moskovsky, C., Assulaimani, T., Racheva, S., & Harkins, J. (2016). The L2 motivational self-system and L2 achievement: A study of Saudi EFL learners. *The Modern Language Journal*, *100*(3), 641–654. https://doi.org/10.1111/modl.12340

Pan, M., Lai, C., & Guo, K. (2025). Effects of GenAI-empowered interactive support on university EFL students' self-regulated strategy use and engagement in reading. *The Internet and Higher Education*, *65*, Article 100991. https://doi.org/10.1016/j.iheduc.2024.100991

Rostami, F., Yousefi, M. H., & Amini, D. (2021). How shifting from teaching Arabic or Persian to English prompts the professional identity: A thematic study. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, *9*(1), 121-141. https://doi.org/10.22049/jalda.2021.27098.1255

Sarmiento-Campos, N.-V., Lázaro-Guillermo, J. C., Silvera-Alarcón, E.-N., Cuellar-Quispe, S., Huamán-Romaní, Y.-L., Apaza, O. A., & Sorkheh, A. (2022). A look at Vygotsky's Sociocultural Theory (SCT): The effectiveness of scaffolding method on EFL learners' speaking achievement. *Education Research International*, *2022*(1), 3514892. https://doi.org/10.1155/2022/3514892

Taye, T., & Teshome, G. (2025). The efficacy of extensive reading strategies for enhancing reading comprehension among 4th year EFL students at Mizan Tepi University. *Social Sciences & Humanities Open*, *11*, Article 101616. https://doi.org/10.1016/j.ssaho.2025.101616

Teng, L. S., Wei, J., & Zhang, L. J. (2024). Longitudinal interactions of L2 learners' motivations and strategic behavior in strategies-based writing instruction: A self-regulated learning perspective. *AILA Review*, *37*(2), 188–214. https://doi.org/10.1075/aila.24026.ten

Uçak, G., & Kartal, G. (2023). Scaffolding design to increase reading comprehension for learners of English through online strategy training. *E-Learning and Digital Media*, *20*(4), 402–423. https://doi.org/10.1177/20427530221111268

van de Pol, J., van den Boom-Muilenburg, S. N., & van Gog, T. (2021). Exploring the relations between teachers' cue-utilization, monitoring and regulation of students' text learning. *Metacognition and Learning*, *16*(3), 769–799. https://doi.org/10.1007/s11409-021-09268-6

Van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review*, *22*, 271–296. https://doi.org/10.1007/s10648-010-9127-6

Wang, J., & Fan, W. (2025). The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: Insights from a meta-analysis. *Humanities and Social Sciences Communications*, *12*, Article 621. https://doi.org/10.1057/s41599-025-04787-y

Wang, Y.-H. (2017). Integrating self-paced mobile learning into language instruction: Impact on reading comprehension and learner satisfaction. *Interactive Learning Environments*, *25*(3), 397–411. https://doi.org/10.1080/10494820.2015.1131170

Wang, Y., Xin, Y., & Chen, L. (2024). Navigating the emotional landscape: Insights into resilience, engagement, and burnout among Chinese high school English as a foreign language learners. *Learning and Motivation*, *86*, Article 101978. https://doi.org/https://doi.org/10.1016/j.lmot.2024.101978

Wang, Y., Derakhshan, A., & Rahimpour, H. (2024). Developing resilience among Chinese and Iranian EFL teachers: A multi-dimensional cross-cultural study. *Journal of Multilingual and Multicultural Development*, *45*(6), 2111–2128. https://doi.org/10.1080/01434632.2022.2042540

Yang, Wei, & Liu. (2025). Intersections between cognitive-emotion regulation, critical thinking and academic resilience with academic motivation and autonomy in EFL learners: Contributions of AI-mediated learning environments. *British Educational Research Journal*, Online Publication. https://doi.org/10.1002/berj.4140

You, C., Dörnyei, Z., & Csizér, K. (2016). Motivation, vision, and gender: A survey of learners of English in China. *Language Learning*, *66*(1), 94–123. https://doi.org/10.1111/lang.12140

Zhang, H. (2024). Translanguaging space and classroom climate created by teacher's emotional scaffolding and students' emotional curves about EFL learning. *International Journal of Multilingualism*, *21*(1), 298–324. https://doi.org/10.1080/14790718.2021.2011893

Zhang, R., Zou, D., & Cheng, G. (2025). Self-regulated digital game-based vocabulary learning: motivation, application of self-regulated learning strategies, EFL vocabulary knowledge development, and their interplay. *Computer Assisted Language Learning*, *38*(8), 1961–2003. https://doi.org/10.1080/09588221.2024.2344555

Zhang, Y., Chen, Q., & Peng, Y. (2025). Guiding the future: The role of L2 self-guides in predicting pre-service language teachers' professional commitment. *International Journal of Applied Linguistics, 35*(4), 2125-2137. https://doi.org/10.1111/ijal.12756

Zhi, R., & Derakhshan, A. (2024). Modelling the interplay between resilience, emotion regulation and psychological well-being among Chinese English language teachers: The mediating role of self-efficacy beliefs. *European Journal of Education: Research, Development and Policy,59*(3), Article e12643. https://doi.org/10.1111/ejed.12643

Authors' Biography

Komeil Jafari Sanksarak received his B.A. and M.A. degrees in Teaching English as a Foreign Language (TEFL) from Islamic Azad University, Sari Branch. He completed his M.A. in 2020 and is currently a Ph.D. candidate in TEFL at Islamic Azad University, Qaemshahr Branch, Iran. His research interests include English language teaching, teacher education, and the development of language skills.

Mehrshad Ahmadian is an assistant professor in TEFL at the Islamic Azad University, Qaemshahr Branch, Iran. He has published papers in various national and international journals. His research domains are (critical) language teacher education, language assessment and teaching all four language skills.

Ebrahim Fakhri is an assistant professor at the Islamic Azad University. He has publications in Springer, Wiley-Blackwell, *Elsevier,* Francis & Taylor, Routledge, and Cambridge University Press. His research domains are bilingualism, listening comprehension, metacognition, EFL motivation, and listening in relation to all other language skills.

Volume 14, Issue 1
Winter and Spring, 2026
pp. 177-198

Journal of Applied Linguistics and Applied Literature: Dynamics and Advances

Technostress Dynamics in EFL Teachers: First- and Second-Order SEM on Computer Self-Efficacy, TPACK, Burnout, Academic Resilience, Support, and Demographics

Hoda Divsar*

**Corresponding Author: Associate Professor of ELT, Department of English Language and Literature, Faculty of Literature and Humanities, University of Guilan, Rasht, Iran, ORCID: 0000-0003-3978-9283*
Email: hodadivsar@guilan.ac.ir

Abstract

Despite the promising role of technology in education, teachers' techno-stress anchors integrating technology into the instructional process. The present research aimed to test a structural model of teachers' techno-stress, technological pedagogical content knowledge (TPACK), computer self-efficacy, environmental support, academic resilience, and burnout. To achieve this goal, 232 EFL teachers completed the Computer Self-Efficacy Scale, Teachers' Technostress Level Defining, TPACK survey, Overall Support questionnaire, Connor–Davidson Resilience Scale, and Maslach Burnout Inventory (2016). Exploratory and confirmatory factor analyses were conducted on all questionnaires, confirming satisfactory validity and reliability indices that supported the subfactors within the constructs. Structural equation modeling (SEM) indicated that: (a) EFL teachers' TPACK, and resilience negatively predicted their technostress while burnout predicted technostress positively; (b) administrative support and computer self-efficacy negatively predicted EFL teachers' technostress through the mediating role of TPACK and resilience; (c) Teachers' burnout serves as a mediator between administrative support and technostress, as well as computer self-efficacy, exhibiting a negative association with latent exogenous variables and demonstrating significant predictive impact on the latent endogenous. (d) Additionally, TPACK and resilience indirectly affected teachers' burnout, demonstrating that elevated levels of TPACK and resilience are associated with reduced burnout and technostress; (e) Technostress was positively associated with age, whereas its relationship with gender was nonsignificant. The analysis yielded instructional recommendations for EFL instructors, educators, and policymakers.

Keywords: technostress, technological pedagogical content knowledge, environmental support, computer self-efficacy, demographic variables

ARTICLE INFO

Research Article
Received: Monday, October 27, 2025
Accepted: Sunday, March 1, 2026
Published: Wedensday, April 1, 2026
Available Online: Sunday, March 1, 2026
DOI: https://doi.org/10.22049/jalda.2025.30073.1734
Online ISSN: 2821-0204; Print ISSN: 28208986

Introduction

Despite the positive aspects of the advancement of hyper-connected technology, the anxiety of dealing with and coping with new information technologies leads to a psychological embarrassment known as technostress. Technostress, therefore, refers to the adverse emotions, thoughts, and attitudes that individuals experience when confronted with new technologies (Mun˜oz et al., 2017). In the academic context, this modern adaptation disorder becomes one of the problems that stands in the way of teachers. Rapid and constant technological developments affect the educational and teaching dynamics in all professional disciplines (Coklar et al., 2017), making it essential for teachers to incessantly expand both their technological and pedagogical knowledge. Therefore, finding out the factors that affect teachers' technostress and overcoming their self-inadequacy to adapt to computer technologies becomes prominent. To stay current with emerging technologies, technological and pedagogical knowledge play a role in decreasing teachers' stress (Zhao et al., 2022). Schmidt et al (2009) defined TPACK as the ability to utilize technology to cover the methods that are used in academic settings with technology, to support students in solving issues they face while learning, to consolidate current knowledge, and to work on technology in terms of keeping and obtaining more up-to-date data and knowledge. Other factors, such as computer self-efficacy, also contribute to practical solutions that can help teachers navigate this anxiety-inducing context (Daniilidou et al., 2020). The belief in one's skills and knowledge of computers, known as computer self-efficacy, empowers teachers to endure setbacks and face complications with minimum tension (Dong, 2020). Accompanied by unceasing support, teachers feel more confident in dealing with evolving technology and with the accompanying novelties in pedagogy. Keeping teachers abreast of the recent technological developments and the pertinent methodologies, supporting them in getting along with technology, and providing backup in times of any breakdown helps them cope with the rapid changes of the information age.

Research on the implications of technostress has demonstrated that technology can evoke unpleasant feelings, such as inadequacy, mental exhaustion, and anxiety, which may lead to burnout (Daniilidou et al., 2020). Burnout denotes the state of mental, emotional, and physical exhaustion that can emerge because of prolonged exposure to emotionally demanding situations, technology hassles, academic adversities, and academic setbacks (Maslach et al., 2016). In a century of stress and tension, burnout might occur due to the misfit between teachers' inadequacy of technological resources and their job demands.

Whatever the reasons for the technostress are, teachers' inner status assists them in managing technostress, diminishes the amount of stress, and accordingly lessens the level of emotional exhaustion. Resilient teachers are less prone to distress in challenging situations, effectively navigating academic challenges and setbacks. They focus on their emotions, strengths, and capabilities rather than on negative aspects and stressors (Joo et al., 2016). What led to the exploration of the relationship among these variables was their contributory effects, since they help teachers tackle their techno-related stresses and setbacks. Little has been done to unveil the causal modeling of the relationship among these variables. It is noteworthy to acknowledge that, except for a few studies, there has been a scarcity of research examining the function of TPACK, resilience, and burnout as mediators, along with environmental support and computer self-efficacy, concerning teachers' techno-stress within a structural model. Thus, this research was executed using the SEM framework to identify the causal relationships among potential contributing factors, namely, teachers' techno-stress, TPACK, computer self-efficacy, administrative support, technical support, and burnout. Using path analysis, the study conjointly examined the predictive TPACK, resilience, and burnout as mediators, along with environmental support and computer self-efficacy, concerning teachers' techno-stress within a structural model. Given the particularly important emphasis on technostress for EFL teachers, the possible reciprocal relations among TPACK, resilience, and burnout were explored. The full model is shown in Fig. 1. The study tried to test the following hypotheses:

H1: Computer self-efficacy significantly influences EFL teachers' technostress.

H2: Environmental support significantly impacts EFL teachers' technostress.

H3: TPACK significantly influences teachers' technostress.

H4. Resilience significantly positively affects teachers' technostress.

H5. Burnout significantly affects teachers' technostress.

H6: The impact of computer self-efficacy on technostress is mediated by TPACK.

H7. The relationship between computer self-efficacy and technostress is mediated by resilience.

H8. Burnout intervenes between computer self-efficacy and technostress.

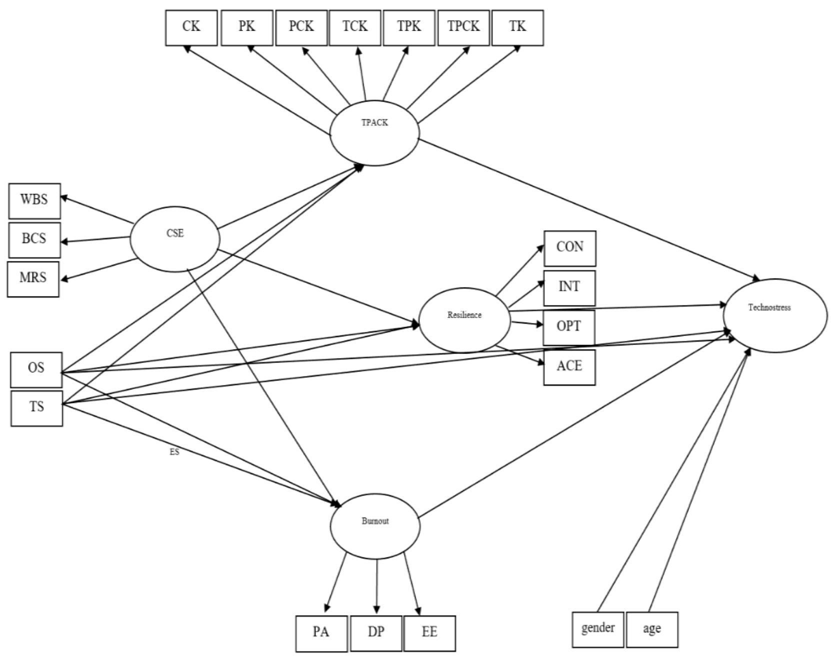
H9. The mediation of TPACK regulates the relationship between environmental factors and technostress.

H10: Resilience mediates between environmental factors and technostress.

H11. Burnout serves as a mediator in the relationship between the environment and technostress.

Figure 1

PLS-SEM Path Model



Review of Literature

Teachers' Technostress

As technology rapidly progresses, technostress has become an essential aspect of individuals' lives, particularly in education, where teachers must not only acquire new technologies but also incorporate them into their teaching to meet various expectations (Özgür, 2020). This tension arises from the gap between contextual constraints and an individual's problem-solving competencies. To keep pace with technological progressions and pedagogical inventions, teachers must continuously boost their knowledge and skills for competent incorporation (Coklar et al., 2017). Research in educational and EFL/TEFL contexts indicates that technostress and technostress creators (e.g., techno-overload, techno-complexity, techno-invasion) unfavorably affect institutional costs (through reduced performance and potential turnover), as well as teachers' job satisfaction and motivation (Aktan & Toraman, 2022; Dong et al., 2020; Zhao et al., 2022).

Technological Pedagogical Content Knowledge (TPACK)

The concept of TPACK merges Technological Knowledge (TK), Pedagogical Knowledge (PK), and Content Knowledge (CK), empowering teachers to adeptly integrate educational technology into the curriculum enterprise (Schmidt et al., 2009). They highlight that TPACK is vital for efficacious technology integration, yet insufficiencies in this knowledge can deter educational outcomes. While previous studies highlight TPACK's role in reducing technostress (Schmidt et al., 2009; Sulla et al., 2022), recent research indicates that knowledge alone does not guarantee successful implementation (Dong et al., 2020; Zhao et al., 2022). The studies conducted in this realm varied from examining the factors that bring about technostress in teachers to the effect of technological knowledge (Joo et al., 2016), to insufficient support (Dong et al., 2020; Özgür,2020), to social and psychological factors (Yeşilyurt et al., 2016). Joo et al (2016) discovered that higher levels of teachers' TPACK are associated with lowered stress related to computer usage, yet their model failed to consider the additional influencing factors. Accordingly, while TPACK reduces stress, its efficacy demands holistic interventions beyond knowledge acquisition.

Computer Self-efficacy

Integrating technology into teachers' pedagogical practices is decisive for promoting educational modernization and growth (Koh et al., 2017). Teachers' confidence in their capacity to exploit technology, known as computer self-efficacy, meaningfully stimulates their efficiency in teaching (Yeşilyurt et al., 2016). Nevertheless, recent scholarship complicates the notion that elevated self-efficacy unilaterally ensures effective technology integration. Studies by Sulla et al. (2022) and Yeşilyurt et al. (2016) reveal that this relationship is dynamically mediated by organizational climate, workload intensity, and teachers' emotional regulation capacities. Furthermore, although higher self-efficacy is positively associated with TPACK and professional resilience, these links are profoundly context-dependent, varying across institutional cultures and support structures (Li, 2023; Skaalvik & Skaalvik, 2010). Accordingly, computer self-efficacy should be reconceptualized not as a static individual disposition but as a socially constructed competence contingent upon systemic support, reflective professional development, and sustainable institutional scaffolding that collectively mitigate technostress and professional burnout.

Environmental Support

The school environment is one of the key conditions for motivating teachers to adopt technology in their instruction (Özgür,2020). Access to technological and administrative support, as well as cooperation among colleagues, strongly improves teachers' intentions to learn technology (Krishnan, 2017). Administrative support, in the form of technical help as well as aid from colleagues, enhances teachers' technology proficiency and self-confidence (Joo et al., 2016). Moreover, technical help offers assurance that contributes to teachers' propensity to use technology (Dong et al., 2020). Earlier studies have indicated a positive correlation between such assistance and decreased technostress (Zhao et al., 2022). The relationships among administration support, TPACK, technostress, resilience, and burnout, nevertheless, still need much scrutiny; thus, the present study aspires to fill the gap in the literature by conducting a thorough investigation into the dynamic interplay.

Burnout

Burnout is a continued reaction to ongoing stressors, marked by emotional collapse, depersonalization, and a reduced sense of personal achievement (Maslach et al., 2016). While traditionally linked to high job demands and limited resources, recent studies highlight the evolving nature of burnout in educational settings, emphasizing how rapid technological integration and shifting pedagogical expectations exacerbate these stressors (Skaalvik & Skaalvik, 2010; Zhao et al., 2022). Critically, this suggests that burnout is not merely a consequence of workload but also a failure of institutional support systems to adapt to changing teaching landscapes. The repercussions extend beyond individual teachers, manifesting as heightened anxiety, strained interpersonal relationships, and waning professional commitment, which collectively undermine both teacher well-being and student achievement (Li, 2023). Emerging evidence underscores the pivotal role of teacher resilience and self-efficacy as mediating factors that can buffer against burnout. However, their development requires deliberate integration into teacher training programs, suggesting that systemic interventions, rather than isolated individual efforts, are essential to reduce burnout and promote sustainable teaching careers effectively (Sulla et al., 2022).

Resilience

Vital in mitigating burnout, resilience is defined as the ability to adapt and cope with challenges such as heavy workloads and high-pressure situations (Li, 2023). Resilient teachers maintain a positive attitude, experience lower stress levels, and achieve greater job satisfaction, resulting in improved student outcomes, including enhanced academic performance and reduced behavioral issues (Daniilidou et al., 2020). Key factors that nurture resilience include social support, positive teacher-student relationships, and effective handling strategies (Polat & İskender, 2018). Notably, resilience can be amended through targeted training, making it essential to understand its expansion, especially among teachers in demanding contexts (Fathi & Saeedian, 2020). Framed within positive psychology, resilience empowers teachers to succeed despite hardship, enhancing their sense of agency and professional fulfilment. Resilience mitigates burnout and elevates efficacy (Daniilidou et al., 2020; Li, 2023), yet studies overlook its mechanisms in EFL teaching under technostress despite its trainability (Fathi & Saeedian, 2020; Polat & İskender, 2018). This study fills that gap via first-/second-order SEM modeling resilience's interplay with self-efficacy, TPACK, support, and demographics.

Method

Participants

This study employed purposive sampling of 232 senior high school EFL teachers, with 51% female (n = 118) and 49% male (n = 114), averaging 37.4 years in age (SD = 8.15). Participants' teaching experience ranged from 4 to 29 years, sourced from Tehran and Karaj. Their professional experience varied: less than five years (2.5%), five to ten years (15.9%), ten to fifteen years (34.1%), fifteen to twenty years (30.5%), and over twenty years (17%). Academic credentials revealed that 46% held a BA, 51% an MA, and 4% PhD, with field specializations in TEFL/applied linguistics (52%), English translation (25%), and English literature (23%). The cohort exceeds the minimum N = 200 threshold for stable first- and second-order SEM with multiple latent constructs (Lakens, 2022; Wolf et al., 2013), yielding robust fit indices ($\chi^2/df < 5$, CFI > .90).

Instruments

The following instruments were utilized in this study:

Computer Self-Efficacy Scale

Developed by Teo and Koh (2010), this twelve-item instrument comprises three sub-components: Basic Computer Skills (BCS), Web-Based Skills (WBS), and Media-Related Skills (MRS). The assessment uses a 5-point Likert scale for each statement, with options from (1) *strongly disagree* to (5) *strongly agree*. In this study, the tool demonstrates a reliability score, measured by Cronbach's Alpha, of .77, which meets the minimum thresholds for research instruments in education and social sciences ($\alpha \geq .70$) (Soriano-Alcantara et al., 2025). CFA confirmed the adequacy of the questionnaire and the structural model, yielding the following fit indices: ($x^2/_{df}$ = 1.28; RMR = .03; RMSEA = .03; AGFI = .95; NFI = .98; GFI = .98; CFI = .99; TLI = .98; IFI = .99).

Teachers' Technostress Level Defining Scale (TTLDS)

It is a twenty-eight-item questionnaire developed by Coklar et al (2017) that employs a 5-point Likert scale. It is organized into five distinct sub-components: Learning-Teaching Process Oriented, Profession Oriented, Technical Issue Oriented, Personal Oriented, and Social Oriented. In this investigation, it demonstrated robust reliability, evidenced by a Cronbach's Alpha coefficient of .85, which exceeds the accepted threshold of .70 for research instruments (Nunnally & Bernstein, 1994). Additionally, CFA provided further validation of the questionnaire's adequacy and the structural model it represents ($x^2/_{df}$= 1.28; RMSEA = .03; RMR = .03; AGFI = .95; NFI = .98; GFI = .98; CFI = .99; TLI = .98; IFI = .99).

TPACK questionnaire

Created by Schmidt et al (2009), this instrument consists of 46 items rated on a 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree), comprehensively addressing all seven components of the TPACK framework. In the present study, the instrument demonstrates a high reliability coefficient of .88, which is above the commonly recommended minimum of .70 for research instruments (Nunnally & Bernstein, 1994). CFA further validated the adequacy of the questionnaire and the structural model, yielding the following fit indices: ($x^2/_{df}$= 1.28; RMSEA = .03; RMR = .03; AGFI = .95; NFI = .98; CFI = .99; GFI = .98; TL I = .98; IFI = .99).

The Teacher Technology Questionnaire (TTQ)

Developed by Lowther and Ross (2000), TTQ includes two subscales: Overall Support (OS) and Technical Assistance (TA). Each subscale contains four items rated on a five-point Likert scale, with 1 representing *strong disagreement* and 5 representing *strong agreement*. The OS subscale measures teachers' views regarding the support received from administrators, colleagues, and parents. The latter evaluates perceptions of technical support, resource availability, and problem-solving assistance. In this study, CFA confirmed the adequacy of both subscales, yielding fit indices of $x^2/_{df}$= 1.28; RMSEA = .03; RMR = .03; AGFI = .95; NFI = .98; CFI = .99; GFI = .98; TLI = .98; IFI = .99. The Cronbach's alpha coefficients for the OS and TA dimensions were .84 and .85, respectively, which are higher than the .70 cutoff considered acceptable for scale reliability (Nunnally & Bernstein, 1994).

***Connor–Davidson Resilience Scale* (CD-RISC)-EFL Teacher Survey**

Designed by Connor and Davidson (2003), this 10-item questionnaire employs a 5-point Likert scale, ranging from 1 (*not true at all*) to 5 (*true all the time*), and encompasses three dimensions: tenacity, optimism, and coping style. Tenacity represents teachers' persistence and dedication in language teaching, while optimism reflects their confidence and positive mindset when encountering difficulties, while optimism signifies their confidence and positive mindset when encountering difficulties. The coping style captures EFL teachers' propensity to resist pressures and proactively navigate risks. In this study, the instrument demonstrated strong reliability, with a Cronbach's alpha coefficient of .83, which surpasses the reliability benchmark of .70 suggested for research measures (Nunnally & Bernstein, 1994). CFA further validated fit indices of ($x^2/_{df}$= 1.28; RMSEA = .03; RMR = .03; AGFI = .95; NFI = .98; CFI = .99; GFI = .98; TLI = .98; IFI = .99).

Burnout Inventory- Educator Survey (MBI-ES)

This 22-item questionnaire (Maslach et al., 2016) utilizes a 7-point Likert scale from 0 (never) to 6 (every day). It is structured into three subscales: Personal Achievement, which comprises 8 items; Depersonalization, entail five items; and Emotional Exhaustion, encompasses nine items. In this study, the Cronbach's alpha of .85 indicates high consistency, surpassing the established threshold of .70 for research instruments (Nunnally & Bernstein, 1994). CFA further confirmed the questionnaire's validity and structural model ($x^2/_{df}$= 1.28; RMSEA = .03; RMR = .03; AGFI = .95; NFI = .98; CFI = .99; GFI = .98; TLI = .98; IFI = .99).

Procedure

Data preservation measures were enacted to certify privacy and anonymity, complying with ethical standards that underlined confidentiality, voluntary participation, and informed consent. To enhance item clarity and scale reliability, all six questionnaires were completed and commented on by 30 Iranian EFL teachers, who reported no comprehension issues. Using purposive sampling, teachers' phone numbers were obtained from the Administrative Office of the Department of Education for communication via Telegram, WhatsApp, or Eitaa. The sample size grew as links were shared among colleagues, with clarifications provided at intervals to prevent confusion. An informative post detailing the scale's measurements and guidelines was pinned for easy access, and inquiries were promptly addressed. Electronic scales were created in Google Docs, and data were collected through direct outreach at various institutes and language schools in Tehran. To expedite collection, phone calls were made to nonrespondents. Familiarity with platforms like Shad, Zoom, and Google Meet facilitated access to e-learning resources. Their consistent engagement with these platforms ensured appropriate exposure to technology-related demands relevant to the construct of technostress. Completed scale forms were reviewed meticulously, resulting in the exclusion of 33 respondents for potential misconduct, such as incomplete submissions or inconsistent markings.

Data Analysis

The research model was evaluated through Structural Equation Modeling (SEM) utilizing SmartPLS version 3.2.4, employing a bootstrapping procedure with 1,000 resamples (Ringle et al., 2020) to estimate the causal relationships suggested by both the outer and inner models. This approach is particularly suitable for models incorporating formative indicators within the framework of second-order modeling (Javidkar et al., 2022)

Results

The Results of the Measurement Model

This study adopted a two-phase analytical strategy, initially evaluating the measurement model before examining the hypothesized relationships and the structural nature of both first- and second-order constructs (Schumacker & Lomax, 2004). The Partial Least Squares (PLS) analysis encompasses two key stages: the first requires weighing the measurement model, focusing on reliability and discriminant validity; the second comprises gauging the structural model. Reliability was tested by examining indicator factor loadings, Cronbach’s alpha, composite reliability (CR), and the Variance Inflation Factor (VIF) (see Table 6). Item loadings and internal consistency values exceeding .70 are deemed acceptable (Fornell & Larcker, 1981). Model validity was established through convergent validity measures—outer loadings and the average variance extracted (AVE)—and discriminant validity, validated via cross-loadings and the Fornell–Larcker criterion (Schumacker & Lomax, 2004). As shown in Table 1, all first-order construct item loadings surpassed the .70 benchmark (Hair et al., 2022), with composite reliability values ranging from .889 to .955, exceeding the .70 threshold and indicating strong construct reliability. Furthermore, AVE values above .70 confirm robust convergent validity. For second-order constructs, outer loadings ranged from .711 to .875, AVE values from .72 to .89, and CR reached .930, collectively affirming satisfactory convergent validity across both construct levels.

Table 1

The Results of Cronbach's Alpha, Average Variance Extracted (AVE), and the Composite Reliability of the Constructs

	α	CR	AVE	Number of Items	loadings
CSE	.801	.826	.724	12	.723-.825
WBS	.840	.844	.820	3	.719-.835
BCS	.860	.870	.861	5	.714-.822
MRS	.891	.919	.820	4	.723-.833
ES	.891	.897	.818	8	.739-.829
OS	.890	.890	.815	4	.793-.842
TS	.889	.891	.811	4	.722-.835
TPACK	.890	.851	.876	46	.779-.855
CK	.891	.841	.870	12	.769-.830
PK	.890	.878	.871	7	.729-.825
PCK	.878	.871	.870	4	.720-.820
TCK	.866	.869	.872	4	.728-.824
TPK	.872	.865	.873	5	.721-.835
TPCK	.890	.884	.875	8	.730-.835
TK	.875	.860	.870	7	.749-.815
RE	.851	.855	.844	10	.740-.850
TEN	.850	.850	.844	3	.750-.815
OPT	.849	.850	.844	4	.711-.815
CoS	.843	.848	.846	3	.733-.835
BO	.872	.889	.898	22	.719-.821
PA	.869	.882	.888	8	.769-.865
DEP	.870	.876	.887	5	.726-.826
EX	.871	.879	.881	9	.779-.825
TechnoS	.890	.899	.899	28	.789-.825
LTPO	.885	.889	.882	7	.749-.805
PO1	.880	.889	.886	6	.779-.875
TIO	.881	.886	.887	6	.759-.865
PO 2	.880	.882	.880	5	.739-.826
SO	.882	.882	.884	4	.751-.813

After thoroughly analyzing the factor loadings, reliability of the latent constructs, and convergent validity, the study advanced to evaluating discriminant validity for both first- and second-order constructs. This was accomplished using the Fornell-Larcker criterion alongside the heterotrait-monotrait ratio (HTMT), methods designed to demonstrate distinctiveness between constructs. To Fornell and Larcker (1981), discriminant validity is settled when a construct's variance surpasses the shared variance with other constructs. As shown in Table 2 and Table 3, the square roots of the average variance extracted (AVE) placed on the diagonal surpass all inter-construct correlations, thereby confirming satisfactory discriminant validity among the first-order latent constructs (Hair et al., 2022).

Table 2

Fornell-Larcker Criterion for First-Order Constructs

	WBS	BCS	MRS	OS	TS	CK	PK	PCK	TCK	TPK	TPCK	TK	TEN	OPT	CoS	PA	DP	EE	LTPO	PO	TIO	PO	SO
WBS	0.74																						
BCS	0.35	0.75																					
MRS	0.07	0.15	0.76																				
OS	0.51	0.15	0.10	0.84																			
TS	0.29	0.64	0.43	0.60	0.83																		
CK	0.28	0.54	0.24	0.55	0.62	0.78																	
PK	0.40	0.52	0.25	0.56	0.42	0.34	0.75																
PCK	0.51	0.55	0.44	0.52	0.45	0.43	0.39	0.86															
TCK	0.46	0.54	0.48	0.43	0.40	0.41	0.42	0.39	0.85														
TPK	0.30	0.44	0.35	0.46	0.49	0.40	0.50	0.30	0.39	0.79													
TPC	0.52	0.51	0.41	0.40	0.39	0.38	0.36	0.33	0.28	0.34	0.81												
TK	0.22	0.45	0.46	0.40	0.35	0.41	0.48	0.31	0.45	0.38	0.34	0.76											
TEN	0.22	0.19	0.16	0.17	0.18	0.20	0.15	0.18	0.28	0.41	0.48	0.15	0.76										
OPT	0.32	0.31	0.40	0.30	0.29	0.37	0.43	0.40	0.26	0.19	0.41	0.64	0.12	0.77									
CoS	0.43	0.33	0.31	0.30	0.32	0.36	0.27	0.30	0.22	0.38	0.19	0.14	0.15	0.17	0.80								
PA	0.40	0.30	0.28	0.26	0.27	0.39	0.35	0.41	0.19	0.41	0.38	0.41	0.10	0.45	0.13	0.79							
DP	0.47	0.30	0.28	0.31	0.29	0.46	0.39	0.44	0.19	0.50	0.42	0.43	0.09	0.19	0.13	0.13	0.82						
EE	0.37	0.34	0.30	0.30	0.32	0.43	0.48	0.51	0.20	0.37	0.50	0.19	0.14	0.32	0.14	0.12	0.46	0.84					
LTP	0.41	0.41	0.44	0.44	0.39	0.40	0.50	0.48	0.30	0.27	0.37	0.31	0.19	0.33	0.16	0.17	0.51	0.51	0.77				
PO	0.32	0.33	0.23	0.23	0.30	0.29	0.22	0.20	0.21	0.46	0.33	0.33	0.22	0.23	0.15	0.22	0.33	0.31	0.50	0.88			
TIO	0.50	0.46	0.41	0.36	0.40	0.39	0.50	0.46	0.19	0.31	0.25	0.19	0.20	0.27	0.19	0.26	0.29	0.20	0.31	0.75	0.80		
PO	0.40	0.49	0.42	0.35	0.49	0.48	0.44	0.46	0.25	0.48	0.46	0.30	0.10	0.25	0.33	0.15	0.20	0.17	0.20	0.28	0.77	0.81	
SO	0.43	0.47	0.40	0.27	0.41	0.30	0.32	0.36	0.19	0.20	0.31	0.29	0.14	0.29	0.37	0.18	0.28	0.14	0.16	0.23	0.59	0.59	0.86

Table 3 offers a distinct perspective on the Fornell-Larcker criterion applied to the second-order constructs, including computer self-efficacy, environmental support, TPACK, resilience, burnout, and technostress. It is evident that the square roots of the average variance extracted (AVE), displayed along the diagonal of the matrix, exceed the correlations with other constructs found in the off-diagonal values (Fornell & Larcker, 1981). This demonstrates that both first- and second-order constructs satisfactorily meet the criteria for discriminant validity (Hair et al., 2022). Collectively, the results affirm that all reflective constructs examined possess the necessary properties to ensure robust reliability as well as convergent and discriminant validity.

Table 3

Fornell-Larcker Criterion for Second-Order Construct

	CSE	ES	TPACK	RE	BO	TechnoS
CSE	0.84					
ES	0.35	0.83				
TPACK	0.43	0.28	0.78			
RE	0.40	0.21	0.43	0.75		
BO	0.60	0.41	0.25	0.35	0.77	
TechnoS	0.31	0.19	0.23	0.37	0.32	0.74

Discriminant validity was examined using the heterotrait-monotrait ratio of correlations (HTMT), with a brink of .85 or below considered indicative of adequate distinctiveness between constructs (Henseler et al., 2015). As shown in Table 2, all HTMT values fell well within this acceptable range, confirming that the constructs demonstrate strong discriminant validity. In this study, technostress creators, computer self-efficacy, TPACK, academic resilience, and environmental support were conceptualized as formative second-order constructs. Specifically, technostress creators encompass multiple facets such as learning-teaching processes, professional concerns, technical issues, personal factors, and social dimensions. Similarly, computer self-efficacy captures rudimentary computer skills, technology-related capabilities, and web-based competencies. The TPACK construct integrates diverse knowledge areas, including technology, content, pedagogy, and their intersections. Burnout and environmental support were each represented as a singular indicator reflecting elements like personal achievement, depersonalization, professional performance, overall support, and technical assistance. Following the two-step procedure, the initial phase involved checking multicollinearity among first-order constructs using the variance inflation factor (VIF). According to Hair et al (2022), VIF standards should remain below 5, and as shown in Table 2, nearly all observed values met this criterion. Further confirmation is illustrated in Table 4, where none of the underlying constructs exceeded the VIF threshold, indicating that multicollinearity is not an issue for the second-order constructs.

Volume 14, Issue 1
Winter and Spring, 2026
pp. 155-174

Table 4

Discriminant validity (HTMT0.85)

	WBS	BCS	MRS	OS	TS	CK	PK	PCK	TCK	TPK	TPCK	TK	TEN	OPT	CoS	PA	DP	EE	LT	PO	TIO	PO	SO
WBS	0.83																						
BCS	0.49	0.82																					
MRS	0.65	0.64	0.80																				
OS	0.37	0.47	0.45	0.84																			
TS	0.43	0.34	0.43	0.44	0.83																		
CK	0.24	0.23	0.32	0.37	0.44	0.79																	
PK	0.44	0.51	0.50	0.39	0.37	0.39	0.76																
PCK	0.45	0.32	0.44	0.20	0.39	0.60	0.34	0.87															
TCK	0.46	0.43	0.36	0.40	0.28	0.53	0.22	0.39	0.87														
TPK	0.30	0.33	0.35	0.41	0.22	0.42	0.38	0.30	0.39	0.87													
TPCK	0.26	0.29	0.42	0.39	0.32	0.36	0.33	0.33	0.28	0.34	0.88												
TK	0.22	0.19	0.46	0.29	0.33	0.41	0.30	0.31	0.45	0.38	0.34	0.88											
TEN	0.22	0.29	0.40	0.17	0.26	0.40	0.22	0.18	0.28	0.41	0.48	0.15	0.84										
OPT	0.32	0.30	0.40	0.30	0.29	0.37	0.40	0.40	0.26	0.39	0.41	0.34	0.42	0.84									
CoS	0.43	0.31	0.45	0.2	0.32	0.35	0.27	0.30	0.22	0.38	0.19	0.14	0.45	0.47	0.85								
PA	0.39	0.34	0.48	0.46	0.27	0.39	0.35	0.41	0.19	0.41	0.38	0.31	0.50	0.46	0.28	0.89							
DP	0.47	0.30	0.38	0.31	0.39	0.46	0.35	0.44	0.19	0.50	0.42	0.33	0.09	0.39	0.13	0.63	0.88						
EE	0.33	0.32	0.44	0.35	0.37	0.43	0.36	0.51	0.40	0.47	0.50	0.19	0.48	0.32	0.14	0.42	0.46	0.88					
LTPO	0.41	0.42	0.46	0.27	0.36	0.40	0.35	0.48	0.30	0.27	0.37	0.31	0.39	0.33	0.16	0.47	0.51	0.51	0.89				
PO1	0.32	0.34	0.33	0.23	0.44	0.29	0.22	0.20	0.21	0.46	0.33	0.33	0.49	0.40	0.15	0.40	0.33	0.31	0.49	0.89			
TIO	0.43	0.40	0.41	0.26	0.37	0.39	0.30	0.46	0.19	0.41	0.25	0.19	0.33	0.38	0.19	0.36	0.49	0.62	0.36	0.60	0.89		
PO2	0.42	0.43	0.45	0.35	0.25	0.48	0.24	0.46	0.25	0.48	0.46	0.30	0.58	0.35	0.33	0.41	0.39	0.57	0.49	0.58	0.65	0.88	
SO	0.44	0.21	0.40	0.29	0.29	0.30	0.32	0.36	0.19	0.20	0.31	0.29	0.46	0.40	0.37	0.43	0.61	0.54	0.53	0.61	0.59	0.54	0.86

To evaluate the second-order constructs within the formative model, two additional criteria were carefully examined: indicator weights and collinearity. Indicator weights reflect the degree of contribution each first-order construct makes toward shaping its corresponding second-order construct, with weights above .10 regarded as meaningful. Unlike reflective models, where items are often interchangeable, formative models require special attention to correlations among indicators, as high correlations can distort weight estimations and their significance (Hair et al., 2022). Collinearity among indicator variables was measured using the variance inflation factor (VIF), where values of five or above signal potential multicollinearity issues. As shown in Table 6, all indicator weights reached significance, ranging comfortably between .240 and .393, indicating robust support from first-order constructs in forming the second-order counterparts. Moreover, the VIF standards ranged from 2.30 to 2.97—well below the expiry of 5—demonstrating that multicollinearity was not a concern. Strong correlations were also observed between each second-order construct and its related first-order dimensions, as confirmed in Table 5, which highlights the significant contributions of all first-order elements. Further, construct validity was assessed by examining both nomological and discriminant validity. Nomological validity tests whether the formative constructs behave as theoretically expected within the model, requiring significant relationships with related constructs. Discriminant validity, on the other hand, ensures constructs are sufficiently distinct, with correlations less than .70 signaling adequate separation. Table 8 reveals significant path coefficients linking the second-order constructs—computer self-efficacy (CSE), environmental support (ES), TPACK, burnout (BU), and resilience (RE)—to technostress, confirming nomological validity. The correlation values ranged between .315 and .685, further substantiating discriminant validity. Collectively, these findings affirm that the formative constructs hold sound validity at both the indicator and construct levels.

Table 5

Valuation of the Second-order Constructs

Construct level					
second-order constructs	**first-order constructs**	**Measure**	**Weights**	***t***	**VIF**
CSE	WBS	Formative	0.380	29.13***	1.71
	BCS		0.370	21.97***	
	MRS		0.351	22.34***	
ES	TS	Formative	0.338	22.97***	1.81
	OS		0.364	23.89***	
TPACK	CK	Formative	0.301	36.34***	1.64
	PK		0.268	29.44***	
	PCK		0.258	29.45***	
	TCK		0.367	30.10***	
	TPK		0.393	31.41***	
	TPCK		0.382	32.18***	
	TK		0.341	33.20***	
BO	PA	Formative	0.240	30.46***	1.82
	DP		0.313	29.25***	
	EE		0.354	28.74***	
RE	TEN	Formative	0.310	22.81***	1.76
	OPT		0.312	30.57***	
	CoS		0.355	28.74***	

TechnoS	LTPO	Formative	0.371	32.19***	1.97
	PO1		0.367	27.07***	
	TIO		0.359	31.71***	
	PO2		0.342	32.73***	
	SO		0.344	23.98***	
Gender			-0.34	6.19***	2.28
Age			-0.31	1.61***	2.38
Teaching experiences			-0.22	4.50***	1.20

Results of the Structural Model and Hypothesis Testing

The structural model was evaluated using a bootstrapping procedure with 1,000 randomly drawn subsamples and a significance threshold set at .05 (Hair et al., 2022). This approach enabled the estimation of standardized path coefficients and the assessment of hypothesis significance. The evaluation included several metrics: the coefficient of determination (R^2), which reflects predictive accuracy; effect size (f^2); predictive relevance (Q^2); and the significance of path coefficients. Assessing multicollinearity was also essential to ensure the accuracy and reliability of the parameter estimates, as high collinearity can undermine the model's predictive power.

The R^2 value represents how well the model explains variance in the endogenous constructs, with benchmarks of .75, .50, and .25 considered substantial, moderate, and weak, respectively (Hair et al., 2022). In this study, computer self-efficacy (CSE) and environmental support (ES) together accounted for 68% of the variance in technostress, while the combined influence of CSE, ES, TPACK, resilience (RE), and burnout (BO) explained 72%, indicating a strong predictive model (Ringle et al., 2020). The mediating variables explained approximately 70% of the variance in the exogenous variables. Effect sizes (f^2) were also calculated, with values of .02, .15, and .35 marking small, medium, and large effects, respectively (Ringle et al., 2020). Here, the effect of predicting TPACK from CSE was moderate, while environmental support, burnout, resilience, and TPACK all exhibited large effects on technostress, except for CSE, which had a moderate effect (see Table 5). The Q^2 values confirmed that the model has adequate predictive relevance, with strong predictive power demonstrated for CSE and ES, moderate for TPACK, and a large negative predictive relevance associated with resilience and burnout.

Predictive relevance, assessed via Q^2 values greater than zero, further confirmed the model's robustness (Ringle et al., 2020). The final assessment focused on path coefficients—direct, indirect, and total effects—reported in Table 6 and illustrated in Figure 1. Direct effects revealed that computer self-efficacy exerted the strongest influence on teachers' technostress (β = .51, p < .001), followed by environmental support (β = .24, p < .001). TPACK also significantly influenced technostress (β = .11, p < .01), while resilience and burnout showed significant negative impacts (β = .11, p < .01; β = -.43, p < .01; β = -.24, p < .01, respectively). Accordingly, hypotheses 1 through 5 received empirical support. The indirect effects on technostress, mediated through TPACK, highlighted the importance of computer self-efficacy (β = .05, p < .05) and environmental support (β = .03, p < .05), thus supporting hypotheses 4 and 7. Similar patterns of indirect influence via resilience and burnout further emphasized the key roles of these constructs in shaping teachers' experiences of technostress.

Table 6

Hypothesized Path Coefficients of the Structural Model with Interaction Effect

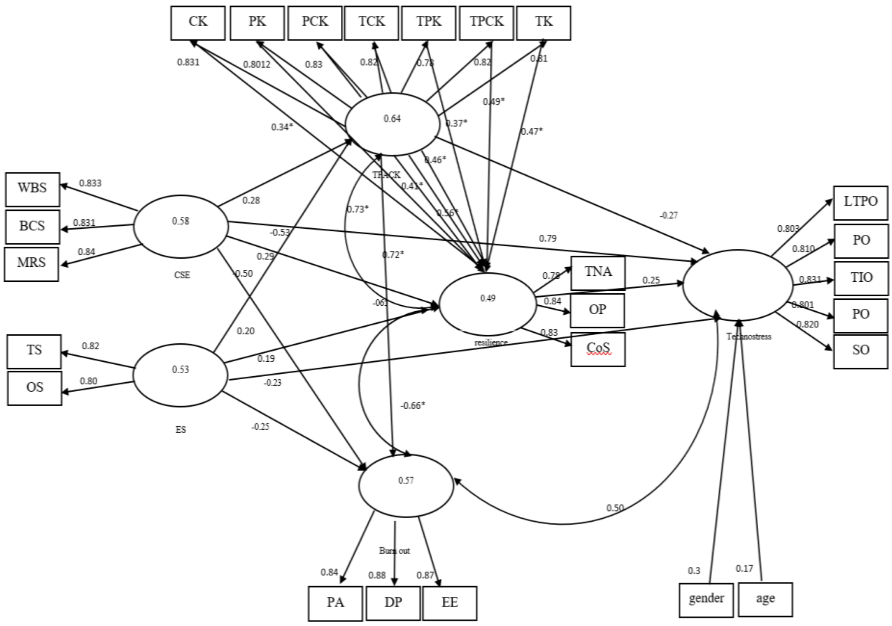
Hypotheses	**Proposed Relationship**	***(β)***	***t***	**Status**	**f²**	**Q²**	**P**	**Bootstrapped 95% CI**	
								LL	UL
		Second-order latent variables							
H1	Computer self-efficacy teachers' → technostress	-0.19	3.45*	supported	0.03	0.01	0.00	-0.44	-0.07
H2	Environmental support → teachers' technostress	-0.14	4.56*	supported	0.01	0.00	0.00	-0.67	-0.18
H3	TPACK → teachers' technostress	-0.18	4.28*	supported	0.31	0.03	0.00	-0.35	0.05
H4	Resilience → teachers' technostress	-0.10	3.75*	supported	0.02	0.00	0.00	-0.26	0.04
H5	Burnout → teachers' technostress	0.13	3.35*	supported	0.05	0.01	0.00	0.07	0.16

H6	Computer-self efficacy → TPACK→ teachers' technostress	-0.30	3.43*	supported	-	-	0.00	-0.26	-0.02
H7	Computer-self efficacy → Resilience→ teachers' technostress	-0.43	4.02*	supported	-	-	0.00	-0.53	-0.19
H8	Computer-self efficacy→ Burn out teachers' technostress	-0.34	3.29*	supported	-	-	0.00	-0.41	-0.17
H9	Environmental support → TPAC K→ teachers' technostress	-0.29	3.45*	supported	-	-	0.00	-0.33	-0.03
H10	Environmental support → Resilience → teachers' technostress	-0.43	4.05*	supported	-	-	0.00	-0.32	-0.17
H11	Environmental support→ Burn out→ teachers' technostress	-0.38	3.22*	supported	-	-	0.00	-0.27	-0.10
	gender→ technostress	0.05	1.12	unsupported	-	-	0.26	-0.04	0.13
	age → technostress	0.22	3.84*	supported	-	-	0.01	-0.31	-0.10

The size of these correlations indicates that the higher the teachers' computer self-efficacy, the less likely they are to experience burnout and technostress. According to Sulla et al (2022), a robust sense of self-efficacy enhances human achievement and personal well-being in various ways, including the capacity to manage stress effectively. In addition, instructors who are provided with support in their careers are more adept at handling diverse stress-related challenges.

Figure 1

Outer Model for PLS-SEM Path Analysis



Discussion

This study investigated how Iranian EFL teachers' TPACK, burnout, and resilience mediate the relationship between computer self-efficacy, environmental support, demographics, and technostress. Results revealed that environmental support significantly and negatively reduces technostress, highlighting the protective role of a supportive, resource-rich atmosphere. Prior research confirms that inadequate support exacerbates technostress (Dong et al., 2019; Joo et al., 2016). Consistent with these results, Joo et al. (2016) underscored the critical role of institutional infrastructure, technical assistance, and collegial collaboration in alleviating technology-induced stress in educational settings. Stakeholders should thus prioritize creating inclusive environments that empower educators and foster sustainable digital engagement. The results indicated that overall and technical supports significantly strengthen teachers' TPACK, with collaborative peer relationships evidencing more impactful than infrastructure or training alone. Prior studies affirm that collegial support and technical assistance ease the burdens of technology integration, boosting computer self-efficacy and TPACK development (Koh et al., 2017; Krishnan, 2017; Yeşilyurt et al., 2016). Strong environmental backing cultivates technological literacy and confidence, encouraging meaningful engagement with educational technology while mitigating technostress. Such support creates a vital ecosystem where educators collaboratively navigate challenges and integrate technology with assurance.

The findings also indicated that both overall and technical support significantly enhance teachers' TPACK, with collaborative peer relationships proving more effective than infrastructure or training alone. Prior studies confirm that collegial support and technical assistance alleviate the challenges of technology integration, thereby boosting computer self-efficacy and TPACK development (Dong et al., 2020; Koh et al., 2017; Yeşilyurt et al., 2016). Strong environmental support fosters technological literacy and confidence, facilitating meaningful engagement with educational technology while reducing technostress. Such support cultivates a vital ecosystem where educators collaboratively address challenges and integrate technology with confidence.

The study found that computer self-efficacy predicts technostress through the mediation of TPACK, indicating that teachers with higher self-efficacy in using computers are more adept at integrating technology into their pedagogical practices. This integration enhances their understanding and application of TPACK, leading to reduced technostress. Consequently, teachers with greater computer self-efficacy are likely to experience lower levels of technostress (Daniilidou et al., 2020; Dong et al., 2020; Fathi & Saeedian, 2020). The results show that the path coefficient from computer self-efficacy to TPACK and technostress is significantly higher than that from environmental support to TPACK, underscoring the greater influence of self-efficacy over school support (Koh et al., 2017; Teo & Koh, 2010; Yeşilyurt et al., 2016). This emphasizes the need for educational leaders to implement training programs that focus on developing teachers' ICT problem-solving skills rather than merely providing procedural guidance for specific technological issues (Mun˜oz et al., 2017). Such strategies can enhance teachers' ICT literacy and confidence in technology use, thereby reducing technostress (Shu et al., 2011).

The relationship between technostress and burnout indicates that teachers experiencing higher levels of technology-related stress are more susceptible to burnout. Educational institutions must acknowledge the impact of technostress on teachers' well-being and implement strategies such as stress management programs and professional development to mitigate burnout. These findings align with those of Zhao et al. (2022) and Califf and Brooks (2020), suggesting that excessive workloads, inadequate training, low computer self-efficacy, and insufficient technical support contribute to teacher burnout. Additionally, the results indicate that technostress intensifies with advancing age, potentially due to prolonged digital exposure and performance demands, mirroring Huffman et al. (2013), who link this to adaptability deficits and cognitive overload in older cohorts. This finding highlights the need for implementing targeted interventions and support frameworks to help older teachers manage technostress more effectively. Conversely, the absence of a statistically significant association between gender and technostress suggests that men and women experience comparable levels of this phenomenon. Aligning with Özgür (2020), who reported that gender did not have a significant influence on technostress, the findings suggest that gender appears to exert a limited influence on technostress, highlighting the need to prioritize age-related determinants in both its theoretical conceptualization and practical management. Furthermore, these results call for future research to examine additional demographic and psychological variables that may shape individuals' susceptibility to technostress. Creating a supportive environment that addresses teachers' technological concerns and provides adequate resources is essential for reducing burnout among EFL teachers (Mun˜oz et al., 2017; Shu et al., 2011).

Conclusion

The study found that age negatively predicts technostress, while gender showed no significant direct relationship. This suggests that younger teachers, regardless of gender, benefit from early exposure to technology, resulting in greater experience and confidence in its use in teaching. Previous research supports the negative correlation between age and technostress (Özgür, 2020; Zhao, 2022). These findings are consistent with various gender-technology studies (Coklar et al., 2017; Krishnan, 2017) but contradict studies suggesting a gender advantage for males (Huffman et al., 2013; Venkatesh & Morris, 2000) or females (Krishnan, 2017). These discrepancies highlight the need to explore underlying gender-specific mechanisms.

This analysis offers several key contributions. First, it demonstrates that higher computer self-efficacy negatively predicts technostress, indicating that greater confidence in technology use can alleviate stress and anxiety. Additionally, superior TPACK negatively predicts burnout, suggesting that teachers proficient in integrating technology, pedagogy, and content knowledge experience reduced emotional exhaustion and depersonalization. These findings underscore the importance of addressing technostress creators, fostering supportive environments, and enhancing teachers' self-efficacy to promote effective technology integration in language education. The findings contribute to the literature on technostress in EFL teaching, offering critical insights for researchers, educators, and policymakers in formulating strategies to support teachers in navigating technological challenges and enhancing their well-being. By mitigating technostress creators, improving computer self-efficacy and TPACK, ensuring adequate environmental support, and addressing burnout, instructors can more efficiently incorporate technology into their instructional practices. Educational institutions and policymakers need to acknowledge these factors and implement appropriate interventions. Future research should investigate collegial support mechanisms to bolster teachers' computer efficacy and TPACK development, as well as explore alternative relationships related to technostress for further insights in the field.

Acknowledgments

The author would like to extend her profound gratitude to the anonymous reviewers whose fruitful comments helped enrich the quality of this manuscript.

References

Aktan, O., & Toraman C. (2022). The relationship between technostress levels and job satisfaction of teachers within the COVID-19 period. *Education Information Technology, 27*(7), 10429-10453. https://doi.10.1007/s10639-022-11027-2

Califf, C., & Brooks, S. (2020). An empirical study of techno-stressors, literacy facilitation, burnout, and turnover intention as experienced by K-12 teachers. *Computers and Education, 1*(57), 103–121. https://doi.org/10.1016/j.compedu.2020.103971

Coklar, A., Efilti, E., & Sahin, L. (2017). Defining teachers' technostress levels: A scale development. *Journal of Education and Practice, 8*(21), 103–115.

Connor, K., & Davidson, J. (2003). Development of a new resilience scale: The Connor-Davidson resilience scale (CD-RISC). *Depression and Anxiety, 1*(18), 76–82. https://doi.org/10.1002/da.10113

Daniilidou, A., Platsidou, M., & Gonida, E. (2020). Primary school teachers' resilience: association with teacher self-efficacy, burnout, and stress. *Electronic Journal of Research in Psychology, 1*(98), 549–582. https://doi.org/10.25115/ejrep.v18i52.3487

Dong, Y., Xu, C., Chai, C., & Zhai, X. (2020). Exploring the structural relationship among teachers' technostress, technological pedagogical content knowledge (TPACK), computer self-efficacy, and school support. *The Asia-Pacific Education Researcher, 29*(3), 147–157. https://doi.org/10.1007/s40299-019-00461-5

Fathi, J., & Saeedian, A. (2020). A structural model of teacher self-efficacy, resilience, and burnout among Iranian EFL teachers, *Iranian Journal of English for Academic Purposes, 9*(2), 2476–3187. https://doi.org/20.1001.1.24763187.2020.9.2.2.1

Fornell, C., & Larcker, D. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. https://doi.org/10.2307/3151312

Hair, J., Hult, G., Ringle, C., & Sarstedt, M. (*2022*). *A primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (3rd ed.). Sage.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Huffman, A. H., Wheelehan, J. D., & Payne, S. M. (2013). The influence of gender roles on technology self-efficacy. *Computers in Human Behavior, 29*(4), 1776–1783. https://doi.org/10.1016/j.chb.2013.02.007

Javidkar, S., Divsar, H., Saeedi, M., & Hadavizadeh, A. (2022). A path analysis of autonomy supportive teaching, EFL Learners' willingness to communicate, self-regulation, academic engagement, and perceived locus of control. *Journal of Modern Research in English Language Studies, 9*(4), 25–49. https://doi.org/10.30479/jmrels.2022.16948.2032

Joo, Y., Lim, K., & Kim, N. (2016). The effects of secondary teachers' technostress on the intention to use technology in South Korea. *Computers & Education, 95*(5), 114–122. https://doi.org/10.1016/j.compedu.2015.12.004

Koh, J., Chai, C., & Lim, W. (2017). Teacher professional development for TPACK-21CL: Effects on teacher ICT integration and student outcomes. *Journal of Educational Computing Research, 55*(2), 172–196. https://doi.org/10.1177/0735633116656848

Krishnan, S. (2017). Personality and espoused cultural differences in technostress creators. *Computers in Human Behavior, 66*(1), 154–167. https://doi.org/10.1016/j.chb.2016.09.039

Lakens, D. (2022). Sample size justification. *Collabra: Psychology, 8*(1), 336–353. https://doi.org/10.1525/collabra.33267

Li, S. (2023). The effect of teacher self-efficacy, teacher resilience, and emotion regulation on teacher burnout: a mediation model. *Frontier in Psychology, 1*(4), 14–22. https://doi.org/10.3389/fpsyg.2023.1185079

Lowther, D., & Ross, S. (2000). *Teacher technology questionnaire (TTQ). Memphis, TN: Center for research in educational policy.* The University of Memphis.

Maslach C., Jackson S., & Leiter M. (2016). *Maslach Burnout Inventory Manual.* (4th ed.). Mind Garden.

Mun˜oz, O., Penalba, F., Sa´nchez, J., & Santos, O. (2017). Reducing techno-anxiety in high school teachers by improving their ICT problem-solving skills. *Behavior & Information Technology, 36*(3), 255–268. https://doi.org/10.1080/0144929X.2016.1221462

Nunnally, J. C., & Bernstein, I. H. (1994). The assessment of reliability. *Psychometric Theory, 3*(3), 248–292.

Özgür, H. (2020). Relationships between teachers' technostress, technological pedagogical content knowledge (TPACK), school support, and demographic variables: A structural equation modeling. *Computers in Human Behavior, 112*, 406–415. https://doi.org/10.1016/j.chb.2020.106468

Polat, D., & İskender, M. (2018). Exploring teachers' resilience in relation to job satisfaction, burnout, organizational commitment, and perception of organizational climate. *International Journal of Psychology, 5*(1), 1–13. https://doi.org/10.17220/ijpes.2018.03.001

Ringle, C., Sarstedt, M., Mitchell, R., & Gudergan, S. (2020). Partial least squares structural equation modeling in HRM research. *International Journal of Human Resource Management, 31*(12), 1617–1643. https://doi.org/10.1080/09585192.2017.1416655

Schmidt, D., Baran, E., Thompson, A., Mishra, P., Koehler, M., & Shin, T. (2009). Technological Pedagogical Content Knowledge (TPACK) The development and validation of an assessment instrument for preservice teachers. *Journal of Research on Technology in Education, 42*(2), 123–149. https://doi.org/10.1080/15391523.2009.10782544

Schumacker, R., & Lomax, R. (2004). *A beginner's guide to structural equation modeling* (2nd ed.). Lawrence Erlbaum Associates Publishers.

Shu, Q., Tu, Q., & Wang, K. (2011). The impact of computer self-efficacy and technology dependence on computer-related technostress: A social cognitive theory perspective. *International Journal of Human-Computer Interaction, 27*(10), 923–939. https://doi.org/10.1080/10447318.2011.555313

Skaalvik, E., & Skaalvik S. (2010). Teacher self-efficacy and teacher burnout: a study of relations. *Teaching Education, 26*(4), 1059–1069. https://doi.org/10.1016/j.tate.2009.11.001

Soriano-Alcantara, J.M., Guillén-Gámez, F.D. & Ruiz-Palmero, J. (2025). Exploring digital competencies: Validation and reliability of an instrument for the educational community and for all educational stages. *Technology Knowledge Learning 30*(1), 307–326. https://doi.org/10.1007/s10758-024-09741-6

Sulla, F., Ragni, B., D'Angelo, M., & Rollo, D. (2022). Teachers' emotions, technostress, and burnout in distance learning during the COVID-19 pandemic. In *Proceedings of Third Workshop on Technology-Enhanced learning environments for Blended Education–The Italian E-Learning Conference, 32*(2), 1–13. https://doi.org/10.3389/fpsyg.2023.1185079

Teo, T., & Koh, J. (2010). Assessing the dimensionality of computer self-efficacy among pre-service teachers in Singapore: A structural equation modelling approach. *International Journal of Education and Development using Information and Communication Technology, 6*(3), 7–18.

Venkatesh, V., & Morris, M. G. (2000). Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage. *MIS Quarterly, 24*(1), 115–139. https://doi.org/10.2307/325098

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement, 73*(6), 913–934. https://doi.org/10.1177/0013164413495237

Yeşilyurt, E., Ulaş, A., & Akan, D. (2016). Teacher self-efficacy, academic self-efficacy, and computer self-efficacy as predictors of attitude toward applying computer-supported education. *Computers in Human Behavior, 64,* 591–601. https://doi.org/10.1016/j.chb.2016.07.038

Zhao, G., Wang, Q., & Wu, L. (2022). Exploring the structural relationship between university support, students' technostress, and burnout in technology-enhanced learning. *Asia-Pacific Education Research, 31*(1), 463–473. https://doi.org/10.1007/s40299-021-00588-4

Author' Biographies

Hoda Divsar is currently an associate professor of ELT at University of Guilan, Rasht, Iran. She has dedicated over 17 years to the field of Teaching English as a Foreign Language (TEFL). Throughout her academic career, she has had the privilege of supervising numerous master's theses and PhD dissertations, guiding emerging scholars in their research endeavors. Her primary areas of interest include Structural Equation Modeling (SEM), critical thinking skills, and curriculum development.

Volume 14, Issue 1
Winter and Spring 2026
pp. 199-218

Journal of Applied Linguistics and Applied Literature: Dynamics and Advances

The Impact of Blended and Virtual Learning on Willingness to Communicate and Oral Proficiency in Iraqi EFL Learners: A Mixed Methods Study

Sami Raihan[1], Saeed Ketabi[2, *], and Zargham Ghabanchi[3]

[1]*PhD Candidate in ELT, University of Isfahan, Isfahan, Iran*
Email: raihansami2018@gmail.com

[2]*Corresponding Author: Associate Professor of Applied Linguistics, Department of English Language and Literature, Faculty of Foreign Languages, University of Isfahan, Isfahan, Iran, ORCID: 0009-0008-7430-7001*
Email: ketabi@fgn.ui.ac.ir

[3]*Associate Professor of Applied Linguistics, Department of English, Faculty of Faculty of Letters and Humanities Ferdowsi University of Mashhad, Mashhad, Iran*
Email: ghabanchi@um.ac.ir; mm12fr@gmail.com

Abstract

This study investigated how blended learning (BL) and virtual learning (VL) impacted Iraqi EFL learners' willingness to communicate (WTC) and their oral proficiency. Utilizing a mixed-methods approach, researchers involved 120 students from Karbala University, who were randomly assigned to face-to-face, virtual, and blended learning groups. Quantitative data were collected through pre- and post-tests, while qualitative insights were gathered from semi-structured interviews with EFL instructors. Findings revealed that both BL and VL significantly enhanced learners' WTC, with no notable differences between the two methods. However, BL demonstrated greater effectiveness in improving oral proficiency compared to VL. The Iraqi EFL teachers generally expressed positive views toward both teaching methods, with a preference for blended learning. The study concluded that while both approaches offered valuable benefits for EFL instruction, the blended learning emerged as a particularly effective method by combining the strengths of virtual and face-to-face teaching. These results could provide insights for developing effective EFL teaching strategies that utilize various learning modalities.

Keywords: blended learning, EFL, oral proficiency, virtual learning, willingness to communicate (WTC), Iraqi EFL Learners

ARTICLE INFO
Research Article
Received: Friday, September 5, 2025
Accepted: Friday, December 19, 2025
Published: Wedensday, April 1, 2026
Available Online: Friday, December 19, 2025
DOI: https://doi.org/ 10.22049/jalda.2026.30974.1875
Online ISSN: 2821-0204; Print ISSN: 28208986

Introduction

The landscape of English as a Foreign Language (EFL) education has undergone significant transformations in recent years, particularly with the integration of technology and the emergence of various instructional modalities (Carter & Darling-Hammond, 2016; Kubota, 2016). This study examines the effectiveness of virtual learning (VL) and blended learning (BL) approaches compared to traditional face-to-face instruction for Iraqi EFL learners, focusing on three key aspects of language learning: vocabulary development, willingness to communicate (WTC), and oral proficiency.

Virtual learning has become increasingly popular as a flexible and accessible substitute for traditional classroom education (Okuto, 2023). This mode of learning allows students to engage with materials at their own speed and take advantage of multimedia tools for interactive experiences (Bahiyah, 2023). Studies suggest that virtual learning can improve vocabulary acquisition by offering a variety of linguistic input and practice opportunities (Hampel, 2010). On the flip side, however, it can also present challenges such as a possible decline in motivation and engagement due to the absence of face-to-face interactions (Singh et al., 2021).

In contrast, in-person learning could be argued to be a foundational component of language instruction, providing direct engagement between educators and students (Singh et al., 2021). This format could promote a nurturing environment for communication, enabling learners to practice speaking and listening skills (Navarrete Piguave & Reyes Borbor, 2023). The social interactions in a traditional classroom can play a significant role in boosting students' willingness to communicate, as immediate feedback and support from both peers and instructors, can foster confidence and motivation (Zhang et al., 2018). Nonetheless, conventional face-to-face methods may not fully address the diverse learning styles and requirements of every student (Lee & Liu, 2024).

Blended learning, which merges aspects of both virtual and in-person instruction, has surfaced as a valuable strategy for EFL education (Wright, 2017). This approach capitalizes on the advantages of traditional and digital learning environments, offering a customized and adaptable educational experience (Alamri et al., 2021). Blended learning promotes vocabulary development through a range of teaching methods and resources while also enhancing oral skills by encouraging interaction in both online and offline environments (Wichadee, 2017). Additionally, the incorporation of technology in blended learning environments can improve students' willingness to communicate by providing platforms that encourage collaboration and peer engagement (Al-Samarraie & Saeed, 2018).

As the demand for English proficiency in Iraq continues to grow, driven by globalization and the need for international communication, there is a pressing need to explore innovative approaches to EFL instruction that can enhance learning outcomes (Kadhim & Kamil, 2024). Traditional teaching methods have often been inadequate in meeting the diverse needs of learners, prompting interest in alternative modalities (Bahiyah, 2023; Okuto, 2023).

Furthermore, the research explores the effects of blended and virtual learning on learners' willingness to communicate and oral proficiency development. The social dynamics of face-to-face learning can influence learners' WTC (Zhang et al., 2018), while the integration of technology in blended learning can enhance learners' willingness to communicate by providing platforms for collaboration and peer interaction (Al-Samarraie & Saeed, 2018).

Additionally, the study seeks to investigate the attitudes and acceptance of Iraqi EFL teachers towards virtual and blended learning contexts. As Blignaut and Els (2010) state, teachers should adopt a 'life-long-learning' attitude to have full engagement with ICT-rich education, and an open attitude is required for the acceptance of technology (Van der Spoel et al., 2020).

By investigating these aspects, the study aims to provide insights into the effectiveness of different instructional modalities in the context of Iraqi EFL education. Understanding how virtual and blended learning environments could impact vocabulary knowledge, willingness to communicate, and oral proficiency can inform the design of more effective language programs (Alamri et al., 2021; Wright, 2017). The findings of this research could possibly address a gap in the literature regarding the efficacy of online and blended language teaching courses compared to traditional face-to-face instruction in Iraq, potentially influencing educational practices, curriculum development, and policy-making in EFL education in the region.

Literature Review

In recent years, higher education institutions have experienced a substantial evolution and are facing novel challenges related to the role and integration of technology in academic settings (Altbach et al., 2010). Enhancements such as the adoption of virtual reality and the integration of blockchain technology are but a few examples emblematic of this transformative revolution. A notable example lies in the immersive capabilities of virtual reality (VR), which have been harnessed for diverse applications such as entertainment, vocational training, and education (Chen & Hsu, 2020; Martín-Gutierrez et al., 2017; Vesisenaho et al., 2019). Virtual Reality (VR) has the potential to support language learners through the provision of immersive contextualization derived from simulated environments, as well as the enhancement of meaning-form connections through the use of multiple sensory inputs. Additionally, it has the potential to enhance the learner's autonomy as suggested by Tseng et al. (2020).

Online education seems to have been increasingly embraced by educational institutes and students because of the various advantages it enjoys. It is not location-bound and can offer the opportunities to the individuals seeking to receive the high-quality education, those who may not have access otherwise (Baranik et al., 2017). The main benefits of virtual classes are their interaction, assistance, and flexibility for the learners who have restricted mobility or lack of educational resources (Arbaugh, 2000). These environments have video conferencing technology that provides face-to-face communication (Manegre & Sabiri, 2022).

Language learners who learn online may experience interruptions from poor internet connections or failing programs. This can certainly be a negative experience for them. However, technological advancements are evolving to decrease these interruptions. While online language learning seems to be influential for L2/FL learning when looking at other subjects, many learners like to take simpler subjects online or through distance learning and take important or difficult subjects face to face (Smith Jaggars, 2014). This general viewpoint may change if more difficult subjects are taught through virtual classrooms where the teacher is more present, given that the course is appropriate to be taught online (Manegre & Sabiri, 2020). To make online language learning effective, it is essential to make sure that both students and teachers have acceptable internet connections, that the challenges are dealt with, and that the students and teachers are guided appropriately in the online courses (Manegre & Sabiri, 2022).

Blended learning seems to have shown promising effects on learners' willingness to communicate (WTC) in various EFL contexts. Tran and Hoa (2024) found that blended learning substantially improved students' classroom engagement in IELTS orientation courses in Vietnam. This increased engagement suggests a potential positive impact on WTC, as students who are more engaged are likely to be more willing to communicate.

Ali et al. (2023) reported that blended learning had a positive impact on first-year Arab university students' oral production and improved students' engagement. While this study did not directly measure WTC, as the improvement in oral production and engagement could be closely related to WTC, the results might be interpreted to suggest a potential positive effect.

Kartal and Balçikanli (2018) explored the changes in Turkish EFL learners' reported willingness to communicate after conducting real-life tasks in a virtual world, which can be considered a form of blended learning. They found that using a virtual world had a positive effect on the reported WTC of participants compared to those who did not use it. These studies suggest that blended learning may have a positive effect on WTC among EFL learners, although more research specifically focused on Iraqi EFL learners is needed.

Virtual learning environments have shown potential in affecting learners' WTC. Lee and Liu (2024) investigated fluctuations in L2 WTC in an online class, finding that EFL learners' levels of L2 WTC are highly dynamic and influenced by various factors in the online environment. This suggests that virtual learning can indeed affect WTC, though the effect may be complex and multifaceted.

Rahimi and Fathi (2024) found that e-tandem, a form of virtual learning, significantly improved EFL learners' WTC compared to conventional classes. This could indicate that certain forms of virtual learning can have a positive effect on WTC. Alqarni (2021) reported higher levels of WTC in online learning contexts compared to face-to-face contexts among Saudi undergraduate EFL learners. This study seems to directly supports the potential of virtual learning to positively affect WTC.

However, Al-Amrani and Harrington (2020) found that some aspects of online social contexts, including limited keyboard skills and lack of immediacy, had negative effects on learners' L2 WTC. This highlights the potential challenges of virtual learning in relation to WTC. Considering the above studies, one may conclude that virtual learning can affect WTC, but the effect may depend on various factors and the specific implementation of virtual learning.

There seems to be limited research directly comparing the effects of blended and virtual learning on WTC. However, some studies provide insights into the relative effectiveness of these approaches. Zhou (2023) compared e-tandem (a form of virtual learning) with conventional in-class speaking tasks and found that the e-tandem group outperformed the conventional group in terms of speaking skills and WTC. Although it did not directly compare blended and virtual learning, the results implicitly suggested that virtual learning could be more effective than traditional in-class methods.

Kartal and Balçikanli (2018) found positive effects of a virtual world (which can be considered a form of blended learning) on WTC. This suggests that blended approaches incorporating virtual elements can be effective in improving WTC. More research is needed to directly compare the effects of blended and virtual learning on WTC, especially in the context of Iraqi EFL learners.

Regarding the relationship between blended learning and proficiency development in ELT contexts, some studies have shown positive effects. Teng and Zeng (2022) found that blended learning had a significant effect on the improvement of oral accuracy and fluency among Chinese middle school students learning English. Ali et al. (2023) reported that blended learning had a positive impact on first-year Arab university students' oral production. Ali et al.'s study could be considered particularly relevant to the present study as it was conducted in an Arab EFL context, which may have similarities to the Iraqi context. Alsowayegh et al. (2019) demonstrated positive effects of using blended learning activities in supporting the improvement of students' learning on listening and speaking at the elementary level. These studies suggest that blended learning can have some positive effects on oral proficiency development, although more research specific to Iraqi EFL learners is needed.

Research has shown mixed but generally positive effects of virtual learning on oral proficiency development. Hoang et al. (2023) found significant improvements in overall oral proficiency, including fluency, coherence, lexical resource, and pronunciation, after a virtual reality intervention with EFL students. Zhou (2023) also found that online language exchanges improved the speaking skills of Chinese postgraduate students in an advanced English program. However, some studies have highlighted challenges. Al-Amrani and Harrington (2020) noted that some aspects of online social contexts could negatively affect learners' communication.

As far as the issue of teachers and students' perspectives on virtual and blended learning is concerned, while there is limited research specifically on Iraqi

EFL teachers' attitudes, some studies provide insights in this regard. Van der Spoel et al. (2020) found that Dutch teachers' perceptions of using technology in their classes changed significantly after experiencing online teaching during the COVID-19 pandemic. This could suggest that teachers' attitudes can evolve with exposure to virtual learning environments.

Ebadi et al. (2020) investigated Iranian university students' attitudes toward a learning management system, finding generally positive perceptions. While this study focused on students, it could provide context for the acceptance of virtual learning platforms in a Middle Eastern setting. Kerimbayev et al. (2020) reported that teachers in Kazakhstan and Slovakia found that using a learning management system provided new ways and tools for instruction, suggesting positive attitudes toward virtual learning tools. Regarding blended learning, Ali et al. (2023) noted that instructors faced barriers such as limited technological infrastructure and uneven digital literacy when implementing blended learning in Arab EFL classrooms. This highlights potential challenges that may influence teachers' attitudes. These studies suggest that teachers' attitudes toward virtual and blended learning can be influenced by their experiences, available infrastructure, and perceived benefits. However, more research is needed specifically on Iraqi EFL teachers' attitudes and acceptance of these learning contexts.

The present study was an attempt to examine the effectiveness of a VL English program in comparison with the same course being presented as BL and Face to Face in the university context. The study was set out to scrutinize the following research questions:

RQ1. Does blended learning have any significant effect on the WTC of Iraqi EFL learners?

RQ2. Does virtual learning have any significant effect on the WTC of Iraqi EFL learners?

RQ3. Is there any significant difference between blended and virtual learning in affecting WTC?

RQ4. Does blended learning have any significant effect on the oral proficiency development of Iraqi EFL learners?

RQ5. Does virtual learning have any significant effect on the oral proficiency development of Iraqi EFL learners?

RQ6. Is there any significant difference between blended and virtual learning in affecting oral proficiency development?

RQ7. What are Iraqi EFL teachers' attitudes toward the VL and BL learning contexts and acceptance of VL and BL English classes?

Method

Participants

The study included 120 Iraqi EFL learners enrolled at Karbala University, Iraq. Participants were randomly assigned to three groups of about 40 students each. The sample included both male and female students.

Instruments

The following instruments were used in the present study:

Oxford Quick Placement Test (OQPT)

The OQPT was used to assess participants' English language proficiency and categorize them into low proficiency (LP) and high proficiency (HP) groups.

Pre-Tests and Post-Tests

IELTS-style examinations were administered before and after the treatment to evaluate participants' vocabulary knowledge.

Semi-Structured Interviews

Interviews were conducted with a sample of 10 EFL instructors to gather qualitative data on attitudes towards VL and BL learning contexts. Questions focused on perceived advantages, challenges, and impacts of BL and VL on engagement, learning outcomes, and motivation.

Design

This study employed a mixed-methods research design with a quasi-experimental approach. Three groups were established:

1. Face-to-face (FACE TO FACE) group (control)
2. Fully online instruction (Virtual Learning - VL) group
3. Blended Learning (BL) group (combination of face-to-face and fully online instruction)

Procedure

Intact English language classes were randomly assigned to three groups: Face-to-Face, VL, and BL. In week one, participants took the OQPT. In week two, vocabulary pre-tests were administered. The treatment lasted one semester (30 sessions) based on the following plan:

- FACE TO FACE group: Traditional classroom instruction
- VL group: Instruction via Zoom platform
- BL group: Combination of face-to-face and online instruction

Vocabulary post-tests were administered after the treatment. In addition, semi-structured interviews were conducted with selected instructors.

Data Analysis

Paired-Sample t-tests and Independent Samples t-test were utilized to analyze data quantitatively. Thematic analysis and content analysis of interview data were used to explore learners' attitudes toward VL and BL learning contexts. All statistical analyses were performed using SPSS version 22.

Results

The First Research Question

The first hypothesis stated: Blended learning does not have an effect on the WTC of Iraqi EFL learners. According to the normality of the variables, a Paired-Sample T-test would be used. First, descriptive statistics are presented and then their variance equality is examined. Table 1 shows the pre-&-Post WTC mean values of the experimental Group 1:

Table 1

Pre-&-Post-WTC Mean Values of Experimental Group 1

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre-test EXP1 WTC	3.017	40	.244	.038
	Post-test EXP1 WTC	3.289	40	.222	.035

In the above table, the pre-and-post-test mean scores are equal to 3.01 and 3.28 respectively in the first experimental group. Table 2 indicates paired-sample T-Test of first experimental group's Pre-&-Post-WTC.

Table 2

Paired-Sample T-Test of First Experimental Group's Pre-&-Post-WTC

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	SD	Std. Error Mean	95% Confidence Interval Lower	Upper			
Pair 1	Pre-&-post-test EXP1 WTC	-.272	.364	.057	-.388	-.155	-4.725	39	.000

According to Table 2, the significance level of the test is 0.000 which is less than 0.05. Therefore, the null hypothesis is not accepted. In other words, blended learning has a significant effect on the WTC of learners. According to the mean scores and upper and lower border of the test, it can be stated that the learners' WTC scores in the post-test are higher than the pre-test. In other words, the BL approach has improved learners' WTC.

The Second Research Question

In this section, the second hypothesis is tested:
H0: Virtual learning does not have a significant effect on the WTC of Iraqi EFL learners.
According to the normality of the variables, a Paired-Sample T-test would be used. First, descriptive statistics are presented and then their variance equality is examined.

Table 3

Pre-&-Post-WTC Mean Values of Experimental 2 Group

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre-test EXP2 WTC	2.966	40	.273	.043
	Post-test EXP2 WTC	3.269	40	.273	.043

In Table 3, the pre-and-post-test mean scores are equal to 2.96 and 3.26 respectively in the first experimental group.

Table 4

Paired-Sample T-Test of Second Experimental Group's Pre-&-Post-WTC

Paired Samples Test									
		Paired Differences					**t**	**df**	**Sig. (2-tailed)**
		Mean	**SD**	**Std. Error Mean**	**95% Confidence Interval of the Difference Lower**	**Upper**			
Pair 1	Pre-&-post-test EXP2 WTC	-.3028	.3437	.054	-.4127	-.1929	-5.572	39	.000

According to Table 4, the significance level of the test is 0.000 which is less than 0.05, and the null hypothesis is not accepted. In other words, virtual learning has a significant effect on WTC of learners. According to the mean scores and upper and lower border of the test, it can be stated that the learners' WTC scores in the post-test are higher than the pre-test. In other words, the VL approach has improved learners' WTC.

The Third Research Question

In this section, the third null hypothesis corresponding to RQ3 is tested: There is no significant difference between blended learning and virtual learning regarding WTC. According to the normality of the variables, an Independent Samples t-test was used. First, descriptive statistics are presented and then their variance equality is examined.

Table 5

The WTC Group Statistics between BL & VL Learning

Group Statistics					
	Exp	N	Mean	Std. Deviation	Std. Error Mean
Post-test WTC	BL	40	3.2898	.22205	.03511
	VL	40	3.2694	.27397	.04332

According to the information of above table, mean scores related to the WTC of learners in the first experimental (BL) and second experimental (VL) groups are equal to 3.28 and 3.26, respectively, in the pre-and-post-tests. One of the presuppositions of the Independent Samples t-test is the homogeneity of variances, for which Levene's Test was used.

Table 6

Independent T-Test of Post-Test WTC

Independent Samples Test											
		Levene's Test		**t-test for Equality of Means**							
		F	**Sig.**	**t**	**df**		**Sig. (2-tailed)**	**Mean Dif.**	**Std. Error Dif.**	**95% Confidence Interval Low**	**Up**
Post-test WTC	Eq. Va. assumed	.045	.833	.365	78	.716	.020	.056	-.090	.13	.13
	Eq. Va. not assumed			.365	74.793		.716	.020	.056	-.09	.13

According to Table 6, the significance level of the test is 0.833, which is larger than 0.05, and the assumption of variance equality can be accepted. Based on the first row of the table and p-value=0.716, which is more than 0.05, it can be stated that there is no significant difference between learners' WTC scores in the first (BL) and second (VL) experimental groups.

The Fourth Research Question

In this section, the fourth null hypothesis corresponding to RQ4 is tested:
H0: Blended learning does not have a significant effect on oral proficiency development.

According to the normality of the variables, the Paired-Sample t-test is used. Table 7 illustrates the pre-&-post-oral mean values of first experimental Group.

Table 7

Pre-&-Post-Oral Mean Values of First Experimental Group

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre-test exp1 oral	21.9250	40	1.96622	.31089
	Post-test exp1 oral	29.6250	40	4.28885	.67813

In Table 7, the pre-and-post-test mean scores are equal to 21.92 and 29.62, respectively, for the first experimental group. Table 8 displays paired-sample T-Test of the first experimental group's pre-&-post-oral proficiency development.

Table 8

Paired-Sample T-Test of First Experimental Group's Pre-&-Post-Oral Proficiency Development

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	SD	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Pre-&-post-test exp.1 oral Prof.	-7.700	5.175	.818	-9.355	-6.045	-9.411	39	.000

According to Table 8, the significance level of the test is 0.000 which is less than 0.05, and the null hypothesis is not accepted. In other words, blended learning has a significant effect on the oral proficiency of learners. According to the mean scores and upper and lower border of the test, it can be stated that the learners' oral proficiency scores in the post-test are higher than the pre-test. In other words, the BL approach has improved learners' oral proficiency.

The Fifth Research Question

In this section, the fifth null hypothesis corresponding to RQ5 is tested: Virtual learning does not have a significant effect on oral proficiency development.

According to the normality of the variables, the Paired-Sample t-test is used. First, descriptive statistics are presented and then their variance equality is examined. Table 9 shows the pre-&-post-oral mean values of second experimental group.

Table 9

Pre-&-Post-Oral Mean Values of Second Experimental Group

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre-test exp2 oral	20.9500	40	2.38532	.37715
	Post-test exp2 oral	26.1000	40	4.90316	.77526

In Table 9, the pre-and-post-test mean scores are equal to 20.95 and 26.10, respectively, in the second experimental group. Table 10 below displays paired-sample T-Test of second experimental group's pre-&-post-oral proficiency.

Table 10

Paired-Sample T-Test of Second Experimental Group's Pre-&-Post-Oral

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	SD	Std. Error Mean	95% Confidence Interval Lower	Upper			
Pair 1	pre-&-post-test exp2 oral	-5.150	5.650	.893	-6.96	-3.343	-5.765	39	.000

According to Table 10, the significance level of the test is 0.000 which is less than 0.05, and the null hypothesis is not accepted. In other words, virtual learning has a significant effect on the oral speaking of learners. According to the mean scores and upper and lower border of the test, it can be stated that the learners' oral speaking scores in the post-test are higher than the pre-test. In other words, the VL approach has improved learners' oral speaking.

The Sixth Research Question

In this section, the sixth null hypothesis corresponding to RQ6 is tested:

H0: There is no significant difference between blended and virtual learning regarding oral proficiency development.

According to the normality of the variables, an Independent Samples t-test is used. First, descriptive statistics are presented and then their variance equality is examined. Table 11 displays the oral group statistics between BL & VL learning

Table 11

The Oral Group Statistics between BL & VL Learning

Group Statistics					
	Exp	N	Mean	Std. Deviation	Std. Error Mean
Oral	BL	40	29.62	4.288	.678
	VL	40	26.10	4.903	.775

According to the information in above table, mean scores related to the oral speaking of learners in the first experimental (BL) and second experimental (VL) groups are equal to 26.10 and 29.62 respectively in the pre-and-post-test. One of the presuppositions of the Independent Samples t-test is the homogeneity of variances, for which Levene's Test was used. Table 12 shows independent T-Test of Post-Test Oral proficiency.

Table 12

Independent T-Test of Post-Test Oral Proficiency

Independent Samples Test										
		Levene's Test		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Dif.	Std. Error Dif.	95% Confidence Interval Lower	Upper
Oral Prof.	Eq. Va. assumed	1.35	.24	3.42	78	.001	3.525	1.029	1.474	5.575
	Eq, Va. not assumed			3.42	76.64	.001	3.525	1.029	1.473	5.576

According Table 12, the significance level of the test is 0.249 which is more than 0.05. Thus, the assumption of variance equality could be accepted. Based on the first row of the table and p-value=0.001, which is less than 0.05, it can be stated that there is a significant difference between learners' oral proficiency scores in the first (BL) and second (VL) experimental groups. According to the mean scores and upper and lower borders of the test, it can be stated that the learners' oral proficiency scores in the first (BL) experimental group are higher than the second (VL) experimental group. In other words, the BL approach has improved learners' oral proficiency.

The Sevenths Research Question

Based on the interview results provided, "What are Iraqi EFL teachers' attitudes toward the VL and BL learning contexts and acceptance of VL and BL English classes?", Iraqi EFL teachers generally showed positive attitudes towards both VL and BL contexts, recognizing their potential benefits while also acknowledging the challenges they present. However, there seems to be a preference for Blended Learning over purely Virtual Learning. Here's a summary of their attitudes:

Attitudes Towards Blended Learning

- Teachers view BL positively, seeing it as a flexible and effective approach that combines the benefits of both face-to-face and online instruction.
- They appreciate BL for its potential to enhance student engagement, provide diverse learning modalities, and offer personalized learning experiences.
- Teachers value the improved communication and immediate feedback opportunities in BL environments.
- They recognize BL's potential to foster collaboration and community building among students.

Attitudes Towards Virtual Learning

- Teachers acknowledge VL's benefits in terms of accessibility, flexibility, and the provision of diverse online resources.
- They appreciate VL's potential to reach a wider audience and its scalability.
- However, they express concerns about reduced interaction, potential isolation, and challenges in maintaining student engagement in fully virtual environments.

Acceptance of BL and VL English classes

- Teachers appear more accepting of BL English classes, viewing them as effective in fostering language acquisition due to the combination of in-person interaction and online resources.

- While they accept VL as a viable option, especially in situations where in-person learning is not possible, they seem to prefer BL for its balanced approach.
- Teachers recognize that the effectiveness of both BL and VL depends on factors such as proper implementation, student needs, and learning objectives.

Challenges and Concerns

- Teachers express concerns about technical issues, the need for self-discipline from students, and potential disparities in access to technology in both BL and VL contexts.
- They highlight the importance of maintaining student engagement and motivation, particularly in VL environments.
- Assessment challenges and ensuring academic integrity in online components are areas of concern for teachers.

In conclusion, Iraqi EFL teachers showed a positive attitude and acceptance towards both BL and VL English classes, with a preference for BL due to its perceived balance between traditional and online learning methods. They recognized the potential of these approaches to enhance language learning while also being mindful of the challenges they present. The teachers' attitudes suggest openness to incorporating technology in language instruction, but with a preference for maintaining some level of face-to-face interaction to support effective language acquisition.

Discussion

The findings of this study provide insights into the effects of blended learning and virtual learning on the willingness to communicate (WTC) and oral proficiency of Iraqi EFL learners. Both BL and VL approaches demonstrated significant positive effects on WTC, aligning with previous research in various EFL contexts. This corroborates the findings of Tran and Hoa (2024), who reported that blended learning substantially improved students' classroom engagement in IELTS orientation courses in Vietnam, suggesting a potential positive impact on WTC.

The improvement in WTC across both modalities may be attributed to reduced anxiety in online and blended environments, increased opportunities for communication through diverse channels, and enhanced preparation time for formulating responses. This is consistent with Alqarni's (2021) study, which reported higher levels of WTC in online learning contexts compared to face-to-face contexts among Saudi undergraduate EFL learners.

When comparing the effects of BL and VL on WTC, no significant difference was found between learners' WTC scores in these approaches. This finding contrasts with the results of Zhou (2023), who found that e-tandem outperformed conventional in-class speaking tasks in terms of speaking skills and

WTC. However, it's important to note that Zhou's study compared virtual learning with traditional in-class methods, not blended learning.

In terms of oral proficiency development, both BL and VL approaches showed significant improvements in learners' oral proficiency skills. This aligns with the findings of Teng and Zeng (2022), who reported that blended learning had a remarkable effect on the improvement of oral accuracy and fluency among Chinese middle school students learning English. Similarly, Hoang et al. (2023) found significant improvements in overall oral proficiency, including fluency, coherence, lexical resource, and pronunciation, after a virtual reality intervention with EFL students.

However, when comparing the effects of BL and VL on oral proficiency development, a significant difference emerged, with BL showing higher scores. This suggests that the combination of face-to-face and online components in BL may provide a more comprehensive environment for developing oral proficiency. This finding is supported by Ali et al. (2023), who reported that blended learning had a positive impact on first-year Arab university students' oral production and improved students' engagement.

The study also explored Iraqi EFL teachers' attitudes toward VL and BL learning contexts and their acceptance of these modalities. Teachers generally recognized the advantages of both approaches, which is consistent with the findings of Van der Spoel et al. (2020), who noted that Dutch teachers' perceptions of using technology in their classes changed significantly after experiencing online teaching during the COVID-19 pandemic. However, our study also revealed challenges similar to those identified by Ali et al. (2023), who noted that instructors faced barriers such as limited technological infrastructure and uneven digital literacy when implementing blended learning in Arab EFL classrooms.

Despite these challenges, the overall attitude appears to be positive, with a willingness to adapt to these new teaching modalities. This aligns with the findings of Kerimbayev et al. (2020), who reported that teachers in Kazakhstan and Slovakia found that using a learning management system provided new ways and tools for instruction, suggesting positive attitudes toward virtual learning tools.

In conclusion, while both BL and VL offer significant benefits for EFL learning, blended learning emerges as a particularly promising approach, combining the advantages of both virtual and face-to-face instruction. This study could contribute to the growing body of evidence supporting the effectiveness of these modern learning environments in enhancing WTC and oral proficiency among EFL learners, while also highlighting the need for further research to optimize their implementation in various educational contexts.

Conclusion

The findings of this study could provide some insights into the effectiveness of blended and virtual learning approaches in EFL contexts, particularly for Iraqi learners. Both BL and VL demonstrate significant potential in enhancing willingness to communicate and oral proficiency development. The comparable effectiveness of BL and VL in improving WTC suggests that the integration of technology and online components plays a crucial role in encouraging learners to engage in communication. This finding could be particularly relevant in contexts where traditional face-to-face instruction may be limited or challenging to implement.

However, the superiority of BL in developing oral proficiency highlights the importance of maintaining some level of face-to-face interaction in language learning. The cooperation between online and in-person components in BL appears to create an optimal environment for oral skill development. Teachers' generally positive attitudes toward these approaches, despite acknowledging challenges, indicate a readiness to embrace innovative teaching methods. This openness is crucial for the successful implementation of BL and VL in EFL contexts.

In conclusion, while both BL and VL offer significant benefits for EFL learning, blended learning emerges as a particularly promising approach, combining the advantages of both virtual and face-to-face instruction. As education continues to evolve, especially in light of global changes and technological advancements, these findings could contribute to providing a foundation for developing effective EFL teaching strategies that leverage the strengths of diverse learning modalities. Future research could further explore the long-term impacts of these approaches and investigate ways to optimize their implementation in various educational contexts.

References

Al-Amrani, S. N., & Harrington, M. (2020). The impact of online social presence on Omani female students' willingness to communicate in English. *Computer-Assisted Language Learning Electronic Journal, 21*(2), 220-237. https://callej.org/index.php/journal/article/view/305

Alamri, H. A., Watson, S., & Watson, W. (2021). Learning technology models that support personalization within blended learning environments in higher education. *TechTrends*, *65*(1), 62-78. https://doi.org/10.1007/s11528-020-00530-3

Ali, S. M., Yunus, K., Alshaikhi, T., Abugohar, M. A., Mohana, T., & Mustafa, T. (2023). The effects of blended learning on first-year Arab university students' oral production. *World Journal of English Language, 13*(8), 146-157.

Alqarni, N. (2021). Language learners' willingness to communicate and speaking anxiety in online versus face-to-face learning contexts. *International Journal of Learning, Teaching and Educational Research*, *20*(11), 57-77. https://doi.org/10.26803/ijlter.20.11.4

Al-samarraie, H., & Saeed, N. (2018) A systematic review of cloud computing tools for collaborative learning: opportunities and challenges to the blended-learning environment. *Computers and Education, 124*, 77-91. https://doi.org/10.1016/j.compedu.2018.05.016

Alsowayegh, N. H., Bardesi, H. J., Garba, I. & Sipra, M. A. (2019). Engaging Students through Blended Learning Activities to Augment Listening and Speaking. *Arab World English Journal (AWEJ) Special Issue on CALL* (5). 267-288. https://dx.doi.org/10.24093/awej/call5.18

Altbach, P.G., Reisberg, L., & Rumbley, L.E. (2010). Trends in Global Higher Education: Tracking an Academic Revolution: A Report Prepared for the UNESCO 2009 World Conference on Higher Education. http://www.unesco.org/tools/fileretrieve/2844977e.pdf

Arbaugh, J. B. (2000). Virtual classroom characteristics and student satisfaction with Internet-based MBA courses. *Journal of Management Education, 24*(1), 32–54. https://doi.org/10.1177/105256290002400104

Bahiyah, N. (2023). Revolutionizing education: Unlocking the potential of asynchronous video for interactive online learning. *International Journal of Education and Humanities*, *3*(2), 187-196. https://doi.org/10.58557/(ijeh).v3i2.149

Baranik, L. E., Wright, N. A., & Reburn, K. L. (2017). Mentoring relationships in online classes, *The Internet and Higher Education, 34,* 65-71. https://doi.org/10.1016/j.iheduc.2017.05.001

Blignaut, A. S., & Els. C. J. (2010). Competency assessment of postgraduate students' readiness for higher education. *The Internet and Higher Education, 13*(3), 101–107. https://doi.org/10.1016/j.iheduc.2010.02.007

Carter, P., & Darling-Hammond, L. (2016). Teaching diverse learners. In Gitomer, D. & Bell, C. *Handbook of research on teaching (5th ed.)* (pp. 593-638). Casemate Group.

Chen, Y., & Hsu, C. (2020). Self-regulated mobile game-based English learning in a virtual reality environment. *Computers and Education, 154*, Article 103910. https://doi.org/10.1016/j.compedu.2020.103910

Ebadi, S., & Ebadijalal, M. (2020). The effect of Google Expeditions virtual reality on EFL learners' willingness to communicate and oral proficiency. *Computer Assisted Language Learning*, Article 1854311. https://doi.org/10.1080/09588221.2020.1854311

Ebadi, S., Khazaei, S., & Bashiri, S. (2020). Technology acceptance of NAVID learning management system in the Iranian medical English courses under the COVID-19 pandemic. *Journal of English Language Teaching and Learning, 12*(26), 401-433. https://doi.org/10.22034/elt.2021.42131.2296

Hampel, R. (2010). Task design for a virtual learning environment in a distance language course. I M. Thomas and H. Reinders, (Eds). *Task-Based Language Learning and Teaching with Technology* (pp.131-153). Continuum Publisher

Hoang, D. T. N., McAlinden, M., & Johnson, N. F. (2023). Extending a learning ecology with virtual reality mobile technology: Oral proficiency outcomes and students' perceptions. *Innovation in Language Learning and Teaching, 17*(3), 491-504. https://doi.org/10.1080/17501229.2022.2070626

Kadhim, S. H., & Kamil, S. A. (2024). The correlation between Iraqi EFL university students' willingness to communicate and productive skills performance. *Alustath Journal for Human and Social Sciences, 62*(42), 22-67.

Kartal, G., & Balçikanli, C. (2018). Willingness to communicate and communication anxiety of EFL learners in a virtual world. *Educational Technology*, *15*(3), 7-25. https://doi.org/10.26634/jet.15.3.14549

Kerimbayev, N., Nurym, N., Akramova, A. & Abdykarimova, S. (2020). Virtual educational environment: Interactive communication using LMS Moodle. *Education and Information Technologies*, *25*(3), 1965-1982. https://doi.org/10.1007/s10639-019-10067-5

Kubota, R. (2016). Neoliberal paradoxes of language learning: xenophobia and international communication. *Journal of Multilingual and Multicultural Development*, *37*(5), 467–480. https://doi.org/10.1080/01434632.2015.1071825

Lee, J. S., & Lee, K. (2019). Affective factors, virtual intercultural experiences, and L2 willingness to communicate in in-class, out-of-class, and digital settings. *Language Teaching Research, 24*(6), 813–833. https://doi.org/10.1177/1362168819831408

Lee, J. S., & Liu, L. (2024). Dynamicity of EFL learners' willingness to communicate in an online class. *Journal of Multilingual and Multicultural Development*, *45*(6), 2078-2096. https://doi.org/10.1080/01434632.2022.2039677

Manegre, M., & Sabiri, K. A. (2022). Online language learning using virtual classrooms: an analysis of teacher perceptions. *Computer Assisted Language Learning*, *35*(5–6), 973–988. https://doi.org/10.1080/09588221.2020.1770290

Martín-Gutierrez, J., Mora, C. E., Anorbe-Díaz, B., & Gonzalez-Marrero, A. (2017). Virtual technologies trends in education. *EURASIA Journal of Mathematics, Science & Technology Education, 13*, 469-486. https://doi.org/10.12973/eurasia.2017.00626a

Okuto, D. A. (2023). *Assessment of Experiences with Virtual Learning Among Students in Universities in Nairobi, Kenya* (Doctoral dissertation, University of Nairobi).

Rahimi, M., & Fathi, J. (2024). Employing e-tandem language learning method to enhance speaking skills and willingness to communicate: The case of EFL learners. *Computer Assisted Language Learning*, *37*(4), 924–960. https://doi.org/10.1080/09588221.2022.2064512

Tang, G. M. (1997). Pocket Electronic Dictionaries for Second Language Learning: Help of Hindrance. *TESL Canada Journal*, *15*(1), 39–57. https://doi.org/10.18806/tesl.v15i1.691

Teng, X., & Zeng, Y. (2022). The effects of blended learning on foreign language learners' oral English competence. *Theory and Practice in Language Studies, 12*(2), 281-291. https://doi.org/10.17507/tpls.1202.09

Tran, T. T., & Hoa, C. T. H. (2024). Impacts of a blended learning model on EFL students' engagement and English proficiency: Evidence from IELTS preparation courses. In Bui, H. P., Kumar, R., & Kamila, N. K. (eds). *Innovations and applications of technology in language education* (pp. 177-192). Auerbach Publication. https://doi.org/10.1201/9781003473916

Tseng, H., Kuo, Y-C., Kuo, & Y-T. (2020). Internet self-efficacy, self-regulation, and student performance. *African-American Adult Students in Online Learning, 19*(2), 161-180.

van der Spoel, I., Noroozi, O., Schuurink, E., & van Ginkel, S. (2020). Teachers' online teaching expectations and experiences during the Covid19-pandemic in the Netherlands. *European Journal of Teacher Education*, *43*(4), 623–638. https://doi.org/10.1080/02619768.2020.1821185

Vesisenaho, M., Juntunen, M., Hakkinen, P., Poysa-Tarhonen, J., Fagerlund, J., Miakush, I., & Parviainen, T. (2019). Virtual reality in education: Focus on the role of emotions and physiological reactivity. *Journal of Virtual Worlds Research, 12*(1), 1–15. https://doi.org/10.4101/jvwr.v12i1.7329

Wichadee, S. (2017). A development of the blended learning model using Edmodo for maximizing students' oral proficiency and motivation. *International Journal of Emerging Technologies in Learning (Online)*, *12*(2), Article 137. https://doi.org/10.3991/ijet.v12i02.6324

Wright, B. M. (2017). Blended learning: Student perception of face-to-face and online EFL lessons. *Indonesian journal of applied linguistics*, *7*(1), 64-71. https://doi.org/10.17509/ijal.v7i1.6859

Zhang, J., Beckmann, N., & Beckmann, J. F. (2018). To talk or not to talk: A review of situational antecedents of willingness to communicate in the second language classroom. *System*, *72*, 226-239. https://doi.org/10.1016/j.system.2018.01.003

Zhang, L. (2023). Trends in educational technology: Transforming learning globally. *International Journal of Business Management and Visuals, 6*(1), 22-28.

Zhou, A. (2023). Investigating the impact of online language exchanges on second language speaking and willingness to communicate of Chinese EFL learners: a mixed methods study. *Frontiers in Psychology, 14*(1), Article 1177922. https://doi.org/10.3389/fpsyg.2023.1177922

Author's Biography

Saeed Ketabi Holds a Ph.D. in Applied Linguistics from the University of Cambridge, England and currently is an associate professor of Applied Linguistics at the Faculty of Foreign Languages of the University of Isfahan. He has published more than 100 articles and 15 books in TEFL and Applied Linguistics.

Volume 14, Issue 1
Winter and Spring, 2026
pp. 219-246

Journal of Applied Linguistics and Applied Literature: Dynamics and Advances

SEMing the Causal Associations Between Teachers' Instructional Practice, Growth Orientations, and Learners' Vocabulary Development

Mohaddeseh Hahsemi-aghdam[1, *], Sima ModirKhamene[2], and Parviz Alavinia[3]

[1]Corresponding Author: PhD Candidate in ELT, Urmia University, Urmia, Iran
Mohashemi076@gmail.com

[2] Associate Professor of TEFL, Department of English language and literature, Faculty of Literature and Humanities, Urmia University, Urmia, Iran
ORCID: 0000-0002-6758-2469; Email: sima.modirkhamene843@gmail.com

[3] Associate Professor of TEFL, Department of English language and literature, Faculty of Literature and Humanities, Urmia University, Urmia, Iran
Email: pevinia2006@gmail.com

Abstract

This study explored the interplay between teaching practices (relevance, feedback/feedforward, and organization and clarity), and students' vocabulary achievement. Additionally, examining the mediating role of students' growth orientation (GO) in this interaction, we also analyzed how teaching practices (TP) interact with GO. A convenient sample of 1,105 tenth-grade high school students participated in the study by completing surveys including adapted measures of GO, TP, and a vocabulary test. Partial least squares structural equation modeling was used which consequently ended in the optimal fit of the overall research model that ensured the right path in the study. The findings revealed that students' perceptions of TP were closely linked to their GO and vocabulary development. It was also found that GO interacted with students' achievements. Furthermore, the results indicated that GO mediated the causal association between students' perception on TP and students' achievements. These results underscore the critical role of fostering GO among students while highlighting the influence of effective TP on shaping students' mindsets, growth goal-setting behaviors, and academic success. The findings further support the idea that targeted interventions aimed at enhancing GO can lead to improved educational achievements.

Keywords: EFL, growth orientation, vocabulary development, teaching practice

ARTICLE INFO
Research Article
Received: Saturday, August 23, 2025
Accepted: Wedensday, April 1, 2026
Published: Wedensday, April 1, 2026
Available Online: Wedensday, April 1, 2026
DOI: https://doi.org/10 10.22049/jalda.2026.30934.1865
Online ISSN: 2821-0204; Print ISSN: 28208986

Introduction

The acquisition of a foreign language is shaped by a complex interaction between personal, behavioral, and environmental factors, each exerting a profound influence on learners' academic trajectories. Drawing on Bandura's (1991) Social Cognitive Theory, language learning is conceptualized as an achievement of triadic reciprocity, where personal beliefs, behaviors, and contextual elements interact dynamically to inform and transform one another. Within this framework, teaching practice (TP) emerges as a pivotal environmental determinant, with empirical evidence consistently highlighting the central role of educators in guiding learners' socialization and academic development. Not only do specific pedagogical strategies matter, but the broader teaching style characterized by relevance, clarity, organization, and feedback mechanisms also significantly shapes students' vocabulary development in language learning contexts. Alongside environmental influences, personal factors such as growth orientation (GO) have garnered increased scholarly attention for their role in fostering vocabulary development. Recent research underscores the interconnectedness of growth mindsets, self-regulatory goals, and learner achievements, positioning GO as a core construct in educational psychology.

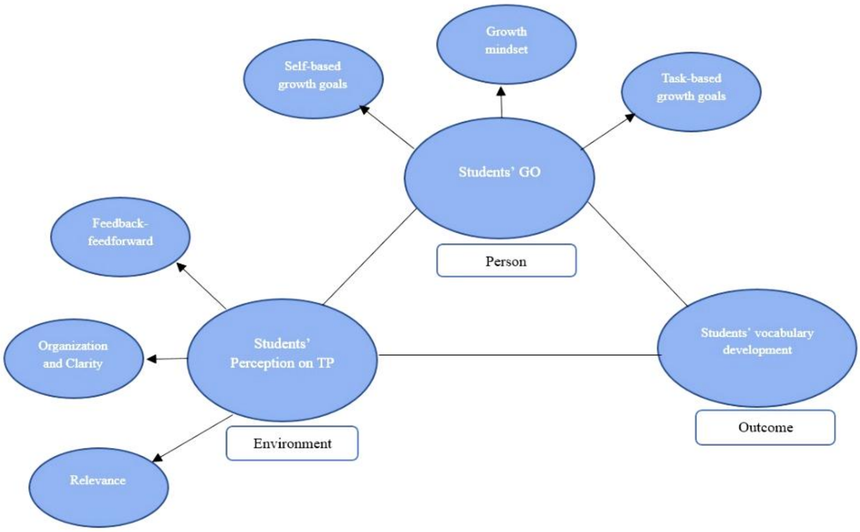
Despite the recognized importance of vocabulary development in educational success, particularly in ESL/EFL settings, the literature reveals a notable gap in understanding how the interplay between TP and learners' GO jointly influences students' achievements (See Figure 1 that shows how the main relationship/s will be examined.

While prior studies have examined the effects of TP and GO on student achievement independently, the integrative examination of the combined impact of these constructs on students' academic (e.g., vocabulary acquisition) achievement remains underexplored. This oversight is especially pronounced in the context of TEFL, where the newly conceptualized unitary construct of GO and its interaction with environmental and behavioral factors have yet to be thoroughly investigated.

The current research contributes to the growing body of research on the determinants of effective language education by systematically examining the interrelations among environmental (TP), personal (GO), and behavioral (vocabulary development) factors within the EFL context. By adopting a comprehensive framework, this research provides nuanced insights into the mechanisms driving student academic success. The findings have the potential to inform targeted interventions and instructional strategies that support the development of well-rounded learners capable of navigating academic challenges. Furthermore, by situating the analysis within the specific domain of TEFL, the study offers novel empirical evidence on domain-specific manifestations of GO, thereby advancing theoretical and practical understanding in the field. To build a comprehensive understanding of this topic, it is essential to examine the existing body of research and scholarly perspectives, which will be discussed in the following review of literature.

Figure 1

Hypothesized Model of Interaction between Different Variables



Review of Literature

To provide a solid foundation for the practical literature review, this study first outlines the relevant terms and theoretical underpinnings.

Theoretical Framework: Bandura's Socio-Cognitive Model

The current investigation is anchored in Bandura's (1986) triadic conceptualization of social–cognitive theory (SCT). This theory provides the foundational lens for understanding GO, teaching practice, and student achievements in language education. SCT posits that human functioning is developed by the continuous interplay of personal factors, behaviors, and environmental influences, with self-efficacy playing a central role in how individuals set goals, persist through challenges, and interpret achievements (Bandura, 1991). This reciprocal determinism is particularly relevant in educational settings, where teachers' beliefs and classroom practices interact with students' cognitive and affective processes to shape learning trajectories.

Growth Orientation

Recent research in psycho-education has been primarily directed towards examining the significance of growth constructs in the academic progression of students (Martin, 2015). Growth constructs, such as growth mindset, self-based growth goals, and task-based growth goals, are manifestations of an individual's dedication to self-improvement. While each of these constructs has its distinct origins and achievements, initial investigations have indicated a certain level of conceptual and empirical commonality among them (Bostwick et al., 2017, 2019; Burns & Martin, 2018). These findings imply that students might inherently possess a fundamental inclination towards growth, a concept identified as GO (Bostwick et al., 2017, 2019).

GO comprises three firmly established growth constructs: growth mindset, self-based growth goals, and task-based growth goals. The concept of growth mindset revolves around individuals' beliefs regarding the changeability of their intelligence, skills, and abilities (Dweck, 2000). On the other hand, self-based growth goals involve learners' aspirations to surpass their previous achievements, focusing on personal progress rather than external comparisons (Martin et al., 2022). Task-based growth goals, however, center around improving performance against predefined task criteria, such as mastery-approach goals that seek to enhance an individual's competence in a particular task or subject area (Elliot & McGregor, 2001). These growth constructs are widely recognized as adaptive educational traits that foster positive academic achievements (Khajavy et al., 2021; Bostwick et al., 2020). Thus, alongside more distinct aspects of other growth constructs, GO captures a complex system of beliefs that benefits students' academic development.

Teaching Practice

Under SCT's triadic model, teaching practice is considered as one of the most pronounced antecedents of goal setting (Burns et al., 2018). On the other hand, teaching factors are found to be of major impact on students' achievements (Martin et al., 2021). Therefore, the study focused on students' perceptions on three key elements of teaching practice (i.e. relevance, organization and clarity, and feedback-feedforward) that are of thought to be of paramount importance in the educational context.

Within our research, the concepts of organization and clarity pertain to the instructor's systematic arrangement and lucidity of the subject matter and assignments, in addition to the effective management of instructional time to enhance learning achievements (Martin et al., 2022). The term feedback-feedforward encompasses corrective information and guidance aimed at improvement provided to students during the learning process (Basso & Belardinelli, 2006). Feedback offers students information about their current task, whereas feedforward provides them with directions for an upswing in their future performances. Together, these practices are essential for fostering self-efficacy, as they help learners recognize their progress and identify actionable steps for continued growth. In the framework of our research, relevance denotes the degree of personal connection and significance associated with the content and tasks presented. This concept builds upon prior investigations and theoretical frameworks that underscore the significance of content and tasks that are easily accessible to learners and that correspond with their existing experiences and knowledge base (Martin & Evans, 2018, 2019). Additionally, it highlights the importance of instructional strategies that involve delivering content and tasks that are both meaningful and pertinent (Lei et al., 2017).

Students' Vocabulary Development

Academic achievement in TEFL are the specific knowledge, skills, and attitudes that learners are expected to acquire or develop as a result of instruction (Brown, 2007). A critical component of academic success in language learning is vocabulary acquisition (Schmidt et al., 2015). This study specifically investigates learners' free productive vocabulary development as an indicator of academic achievement. This refers to the lexicon readily employed in spontaneous written expression as the indicator of academic achievement (Nation, 2001). According to studies (Hajabi et al., 2018; Duran & Kääntä, 2023), the integration of free vocabulary into language instruction can substantially advance the cultivation of communicative competence and encourage more genuine language utilization. Their adaptability permits the formation of innovative terms through processes such as compounding, derivation, and borrowing, thereby mirroring the fluidity of language and its ability to evolve in response to emerging concepts and technologies (Nation & Newton, 1997).

Review of Empirical Studies

A critical examination of the practical literature offers valuable insights into the evolving landscape of research and application within this field. This section provides a comprehensive overview of the existing practical research relevant to the topic, highlighting key findings and identifying gaps that inform the present study. Zeng et al. (2019) identified growth mindset, perseverance, and well-being as direct predictors of work engagement in a sample of 472 Chinese secondary teachers. Their structural equation modeling demonstrated independent predictive capacities for each construct, advocating for targeted professional development programs to cultivate these resilience-related attributes. The findings align with Bostwick et al. (2019), whose longitudinal analysis of students' mathematics achievements established GO as a significant predictor of enhanced academic engagement and achievement.

Expanding this paradigm, Bostwick et al. (2020) employed multilevel modeling to analyze 1,414 students and 91 teachers, revealing positive cross-level associations between mathematics achievements and classroom/teacher GO. While student-level GO consistently predicted engagement and achievement, the study notes potential contextual moderators requiring further exploration, particularly regarding implementation fidelity in diverse educational ecosystems.

Granziera et al. (2022) conducted a cross-national investigation comparing 2,510 Singaporean high schoolers and 119 Australian elementary students. Instrumental (vs. emotional) teacher support exhibited stronger associations with academic buoyancy and engagement across both cohorts, with longitudinal Australian data further linking buoyancy to skill development and aspirational achievements.

In another notable study, Martin et al. (2022) analyzed 61,879 Australian secondary students to demonstrate instructional support's mediated effect on engagement through growth goal setting. Notably, growth goals attenuated socioeconomic and prior achievement disadvantages, with feedback-forward practices and relevance-focused teaching emerging as key pedagogical levers. The study advances growth goal theory by positioning it as both a mediator of instructional quality and a buffer against systemic inequities.

In their seminal work, Bostwick et al. (2022) investigated the relationship between GO and mathematics achievements over a one-year period using a multilevel modeling approach that incorporated cross-level interactions. Their sample comprised 851 adolescents distributed across 71 classrooms, each led by a distinct teacher. The study revealed that GO positively influenced multiple tiers of mathematics achievements, including adaptive and maladaptive engagement as well as academic achievement. Importantly, classroom-level GO moderated the interaction between individual students' GO and their adaptive engagement, suggesting that cultivating growth-oriented environments at both student and classroom levels can enhance mathematics engagement and performance. These findings underscore the potential benefits of targeting GO across educational strata to improve students' achievements.

As the literature suggests, extensive research has established the foundational link between discrete growth constructs and academic progress; however, recent scholarship emphasizes a unifying construct, GO, as a critical determinant of student success. Despite its conceptual significance, GO has yet to be rigorously investigated as an integrated framework. Furthermore, the influence of GO on students' emotional well-being and vocabulary development has been underexplored, limiting comprehensive understanding of its role in fostering positive educational and psychological achievements. This gap underscores the need for research that elucidates the complex mechanisms through which GO operates within learning environments. Instructional practices, particularly those involving relevance, organization, clarity, and feedback-feedforward mechanisms are recognized within Social Cognitive Theory as pivotal antecedents to goal formation. Yet, empirical studies exploring how these pedagogical factors specifically facilitate growth-oriented goal setting and impact student achievements are scarce. Addressing this gap could inform more effective teaching strategies that promote GO and enhance academic engagement. Moreover, the dynamic interplay between GO, TP, and student achievements remains insufficiently understood. Investigating these interactions promises to clarify the pathways through which educational environments shape cognitive dimensions of learning. Finally, the prevailing literature predominantly adopts a domain-general perspective, with limited attention to domain-specific investigations. To date, no study has systematically examined the interrelations among growth constructs, TP, and student achievements within a distinct academic discipline. Filling this gap is crucial for developing targeted interventions that address the unique challenges and opportunities inherent in specific fields of study. Accordingly, the subsequent research questions were formed to guide this investigation:

1) Does TP significantly predict learners' GO?
2) Does learners' GO significantly predict their vocabulary development?
3) Does learners' GO serve as a mediator between TP and students' vocabulary development?

Method

Participants

This study recruited a convenience sample of 1,105 Iranian 10th-grade students (727 females, 378 males; mean age = 15.58 years) from 150 classrooms across four cities in northwest Iran. The sampling frame included both urban and semi-urban public high schools, selected on the basis of teachers' and school principals' willingness to participate and accessibility to the research team.

To achieve a large sample size required for robust statistical analysis, we employed a two-stage convenience sampling strategy. First, school principals were approached for institutional consent. Second, English teachers responsible for 10th-grade classes were invited to administer the student survey in one of their 10th-grade classes, selected at the teacher's discretion.

To contextualize the findings, we collected data on students' prior English achievement (self-reported grades) and language exposure outside school. This information was used to characterize the sample and explore possible confounding variables. Ethical approval as well as written informed consent were obtained from the ministry of education and all students, respectively. Moreover, participation was voluntary, and confidentiality of students' responses was guaranteed.

Instrument Development and Validation

In addition to demographics and background, the student survey consisted of psychometric inventories (28 items across two constructs of TP and GO and measures of vocabulary with 30-items). All instruments, except vocabulary test, were administered in Persian. For measures originally developed in English, a rigorous translation and back-translation process was conducted by two independent bilingual experts, followed by pilot testing that ensured semantic and conceptual equivalence. Prior to the main study, a pilot test involving 30 students, matched to the main sample in age, grade, and school type, was conducted to evaluate the clarity, validity, and reliability of the instruments. Feedback from pilot participants led to minor wording adjustments for clarity and cultural appropriateness. Cronbach's alpha was calculated to measure internal consistency and construct validity was assessed via expert review.

Students' perceptions of teaching practice were assessed using an 18-item questionnaire across three dimensions: instructional relevance, organization and clarity, and feedback-feedforward. The adapted forms of the Teaching Relevance Scale (Martin et al., 2022), Organization and clarity (Blaich et al., 2016), and Feedback-feedforward (Martin & Evans, 2018), were used to evaluate instructional relevance, Organization and Clarity and Feedback/Feedforward. All items were rated on a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree). Cronbach's alpha for the overall TP scale in the pilot was 0.90 (main study: 0.88). Subscale alphas ranged from 0.71 to 0.82. To minimize participant fatigue and ensure survey completion, we selected the highest-loading and most representative items from each subscale, based on expert consensus.

Three indicators of growth mindset, self-referential growth goals, and task-oriented growth goals were measured using a ten-item questionnaire. Growth mindset was assessed using 2 items adapted from Dweck (2000), assessing beliefs about intelligence malleability. Self-based growth goals with 4 items which were adapted from Martin (2006) covered specific, self-improvement, competitive, and challenging goals. Task-oriented growth goals were examined using 4 items adapted from Elliot and McGregor (2001), covering mastery/performance approach and avoidance. The inventory had a 5-point Likert scale format as above. Cronbach's alpha for the overall GO scale in the pilot was 0.72 (main study: 0.74). The growth mindset subscale, with only 2 items, had lower reliability ($\alpha = 0.67$); this is a noted limitation.

Exploratory factor analysis (EFA) was conducted separately for each theoretically derived subscale on the full sample (N = 1105) using principal axis factoring with promax rotation.

The 18-item Teaching Practice (TP) subscale showed excellent sampling adequacy (KMO = .93; Bartlett's test χ^2 (153) = 8431.41, $p < .001$). Parallel analysis supported a two-factor solution accounting for 44.4% of the variance (Factor 1 = 28.2%, Factor 2 = 16.2%). All items loaded $\geq .42$ on their primary factor with minimal cross-loadings (Table 1). Cronbach's α for the full TP subscale was .91, indicating very good internal consistency.

Table 1

Factor Loadings for the TP Subscale (N = 1105)

Item	Factor 1	Factor 2
1	0.423	0.183
2	0.574	0.157
3	0.581	0.214
4	0.559	0.233
5	0.517	0.185
6	0.592	0.186
7	0.577	0.312
8	0.611	0.270
9	0.616	0.329
10	0.585	0.281
11	0.617	0.265
12	0.595	0.143
13	0.584	0.181
14	0.611	0.280

15	0.571	0.328
16	0.232	0.752
17	0.122	0.902
18	0.171	0.804

The 10-item Growth Orientation (GO) subscale also showed good sampling adequacy (KMO = .79; Bartlett's test $\chi^2(45) = 3128.70$, $p < .001$). Parallel analysis clearly indicated a single factor (initial eigenvalues: 3.61, 0.92, 0.84, …), explaining 36.1% of the variance. All items loaded between .44 and .76 on this common factor (Table 2). Internal consistency was acceptable (Cronbach's $\alpha = .78$).

Table 2

Factor Loadings for the GO Subscale (N = 1105)

Item	Factor 1
19	0.69
20	0.62
21	0.76
22	0.68
23	0.64
24	0.59
25	0.52
26	0.44
27	0.56
28	0.61

Vocabulary development was assessed via a 30-item multiple-choice test developed for this study. Items were based on the 10th-grade English curriculum and reviewed by three experienced English teachers for content validity. The test was piloted with 30 students. The internal consistency of the vocabulary test was estimated using the KR-21 on a pilot sample of students. The KR-21 coefficient was 0.90 indicating very good reliability. Item difficulty and discrimination indices were also calculated, and poorly performing items were revised or replaced.

Data were collected over a four-week period in fall 2024. To maximize inclusivity and participation, both paper-based and online (Google Forms) versions of the survey were offered. The researchers endeavored to gather data from participants present at the time through the administration of paper questionnaires, while simultaneously accommodating those participants who exhibited a preference for engaging in the study via the completion of an online questionnaire.

For paper surveys, teachers were instructed to seat students as in examination conditions, allocate 45 minutes for completion, and emphasize the anonymity of responses. For online surveys, teachers distributed the survey link to students and guided students on how to complete it. To minimize data entry errors, paper responses were double-entered and cross-checked. Online data were downloaded directly from Google Forms. Incomplete or duplicate responses were excluded.

Data Analysis

Firstly, descriptive statistics were obtained to get a comprehensive view of data. In addition, prior to performing structural equations, Kolmogorov-Smirnov test was conducted; since the data distribution was not normal, via Smart PLS-3 software, the partial least squares structural equation modeling including measurement model fit, structural model fit, and overall model fit was used to test the proposed hypotheses. This model measures the casual association across students' perception on TP with students' vocabulary achievement considering the mediating role of students' GO.

Measurement of Model fit

The fit of measurement model was tested in terms of divergent validity, indicator reliability, and convergent validity.

Indicator reliability: The results of the variance inflation test were examined in order to assess the reliability of the indicator. Two tests of the PLS algorithm and bootstrapping were used to perform Factor loadings, t-values, and reliability of the determined components of the research model (see Table 3).

Table 3

Factor Loadings, T-Values, and Reliability of the Determined Components of the Research Model

Variables	Sub-scales	Item	Factor Load	t Value	VIF
Students' Perception on TP	Feedback-Feed forward	1	.621	22.565	1.264
		2	.767	50.618	1.567
		3	.737	46.816	1.462
		4	.773	48.174	1.572
		5	.674	30.929	1.310
	Organization and Clarity	29	.492	2.682	1.347
		30	.433	3.131	1.303
		31	.430	2.854	1.194
		32	.473	3.118	1.214
		33	.446	2.879	1.363
		34	.411	2.988	1.228
		35	.481	3.337	1.199
		36	.434	3.115	1.333
		37	.485	3.040	1.214
		38	.487	2.998	1.205

	Relevance	39	.815	57.076	1.444
		40	.843	69.742	1.567
		41	.706	30.465	1.245
Students' Go	Growth Mindset	6	.900	3.292	1.067
		7	.927	2.645	1.067
	Self-based growth goals	8	.519	17.910	1.056
		9	.781	49.643	1.353
		10	.734	37.365	1.281
		11	.588	22.740	1.112
	Task-based growth goals	12	.663	25.518	1.090
		13	.616	22.013	1.058
		14	.418	10.426	1.035
		15	.633	19.394	1.081
Students' vocabulary development	Students' vocabulary development	1	.450	3.695	1.182
		2	.476	2.675	1.180
		3	.476	3.089	1.107
		4	.441	3.310	1.232
		5	.451	4.565	1.398
		6	.580	6.368	1.500
		7	.402	2.783	1.279
		8	.499	5.309	1.090
		9	.484	3.403	1.368
		10	.584	6.591	1.483
		11	.498	3.118	1.120
		12	.425	3.395	1.184
		13	.537	5.420	1.414
		14	.469	3.368	1.351
		15	.428	3.339	1.198
		16	.403	2.996	1.170
		17	.467	4.955	1.142
		18	.470	2.747	1.164
		19	.484	2.944	1.159
		20	.433	4.001	1.569
		21	.468	4.294	1.508
		22	.578	7.258	1.391
		23	.446	4.185	1.144
		24	.471	2.702	1.185
		25	.479	3.501	1.241
		26	.497	5.121	1.439
		27	.429	3.292	1.646
		28	.465	5.599	1.302
		29	.536	5.626	1.521
		30	.475	7.350	1.288

According to Table 3, the absolute value of the standard factor loading for all items is higher than 0.4; also, the absolute value of their t-statistic is higher than 2.58. Accordingly, all items for the measurement of the research model have sufficient validity. Moreover, since the VIF value for the all items is not higher than 5, there is no problem of multiple collinearity for testing the first research model.

Convergent validity test. The results of composite reliability indices, Cronbach's alpha coefficient, and average variance extracted (AVE) that are used to examine convergence are demonstrated in Table 4.

Table 4

Convergent Validity Test for the Model

Variable	Composite Reliability	Cronbach's Alpha	AVE
Feedback-Feed forward	0.761	0.840	0.514
Organization and Clarity	0.716	0.710	0.567
Relevance	0.797	0.832	0.625
Growth Mindset	0.713	0.701	0.525
Self-based growth goals	0.755	0.762	0.541
Task-based growth goals	0.765	0.775	0.548
Students' vocabulary development	0.740	0.785	0.549

As summarized in Table 4, the obtained values for composite reliability, Cronbach's alpha and for all research variables were greater than 0.7. Also, the mean variance extracted for all constructs (variables) was greater than 0.5; the constructs of the model thus have sufficient validity in terms of convergence and correlation.

Divergent Validity Test. Finally, the discriminant validity index was used to test the validity of the constructs (Table 5). The values on the main diagonal of this matrix represent the square root of the AVE. Table 5 shows that the values on the main diameter have the highest number of columns, which indicates the appropriate validity of the constructs. According to the results of this section, the research model is confirmed in terms of measurement. Next, the first research model is examined in terms of structure.

Table 5

Divergent Validity Test for the Model

Variable	1	2	3	4	5	6	7	8	9
1. Organization and Clarity	0.665								
2. Feedback-feed-forward	0.456	0.717	0.845						
3. Growth Mindset	0.444	0.664	0.677	0.725					
4. Relevance	0.386	0.421	0.577	0.578	0.790	0.813			
5. Self-based growth	0.476	0.331	0.480	0.288	0.376	0.664	0.822		
6. students' vocabulary development	0.175	0.139	0.195	0.204	0.235	0.529	0.386	0.537	
7. Task-based growth	0.314	0.257	0.300	0.234	0.254	0.235	0.354	0.590	0.796

Structural Model Fitting

Structural fit indices include the coefficient of determination or R^2 (i.e. a measure that indicates the effect of an exogenous variable on an endogenous variable), the Q^2 (i.e. or the Stone Geisser index determines the predictive power of the model) criterion, and the t value. The blindfolding test was used to perform this fit (Table 6*)*. Since R^2 and Q^2 are above 0.19 and 0.15, respectively, therefore the structural fit of the model can be trusted.

Table 6

Results of Structural Fit of the First Research Model

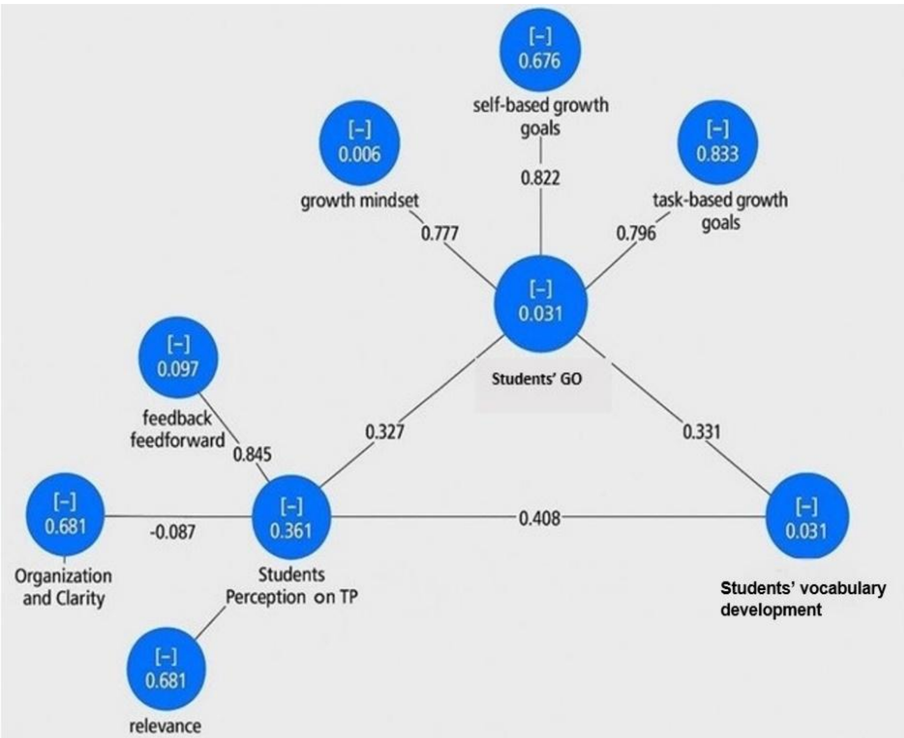
Variable	Communality (Q^2)	Redundancy (R^2)
Feedback-Feed forward	0.475	Exogenous variable
Organization and Clarity	0.225	Exogenous variable
Relevance	0.465	Exogenous variable
Growth Mindset	0.415	0.502
Self-based growth goals	0.497	0.594
Task-based growth goals	0.455	0.516
Students' vocabulary development	0.488	0.301

Overall Model Fit

In order to examine the overall model fit, the GOF criterion proposed by Tenenhaus et al. (2004) was calculated. Three values of 0.01, 0.25, and 0.36 are introduced as weak, medium, and strong values for GOF. According to the formula, the GOF criterion was 0.407. Based on the aforementioned classification, the number indicates the optimal fit of the overall first research model. The high GOF criterion confirms the overall model with power. Therefore, it is concluded that the research model has sufficient validity in general. The results of the research hypothesis test related to the model were carried out using the bootstrap test, the results of which are presented in Figure *2* (model tested based on path coefficients) and Figure 3 (model tested based on t-statistics).

Figure 2

The Model tested based on path coefficients



In this study, the coefficient of determination (R^2) was used to determine the validity of the model. This coefficient of explanatory variance measures an endogenous variable by exogenous variables., The coefficient of determination for the variable "GO" is 0.107 (Figure 2). This means that about 10 percent of the changes related to this variable are explained by changes in the variable of students' perception on TP. The coefficient of determination for the variable "student achievements" is 0.364. This means that 36.4 percent of the changes related to this variable are explained by changes in the variables of GO and students' perception on TP. The tested research model is also presented based on the t-statistic.

Figure 3

The Model Tested Based on T-Statistics

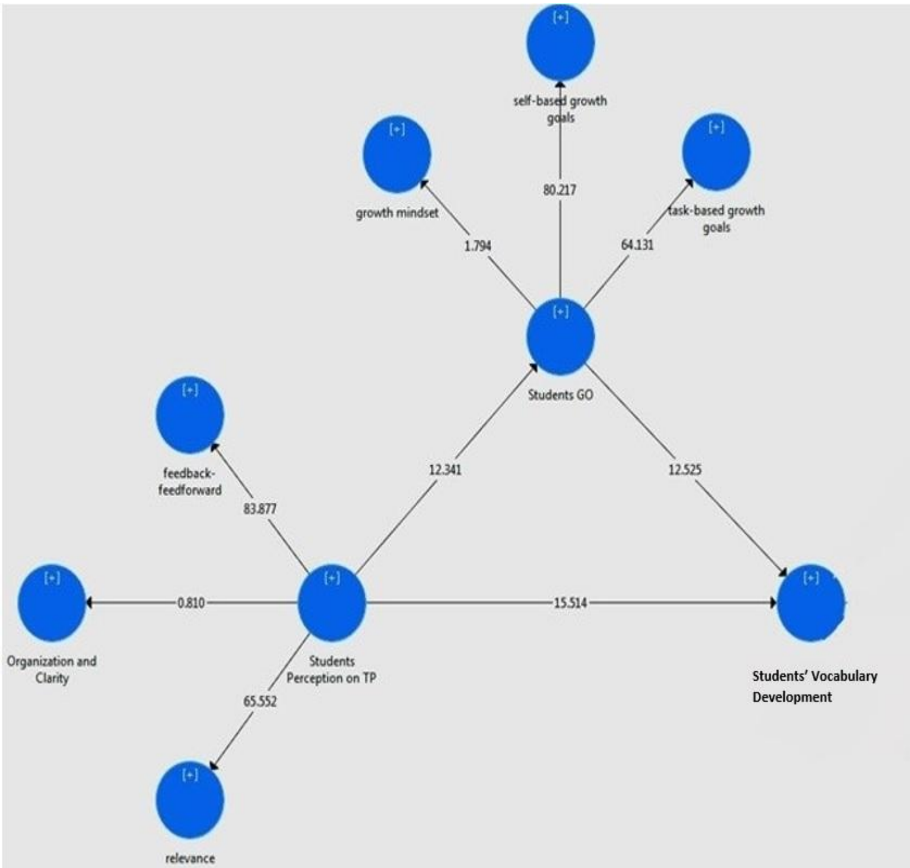


Figure 3 shows the significance of the relationship between the constructs based on the t-statistic. For this purpose, a bootstrapping test was used. According to this figure, if the t-statistic values are above 1.96, the relationship is significant.

Results

The results of the analysis depicting the causal relationship between students' perceptions of TP and GO are presented in detail in Table 7.

Table 7

The Path Coefficient and T-Statistic for the Casual Association across Students' Perception on TP and Growth Orientation

	The path coefficient (β)	t Value	Result
Feedback-feed Forward ↔ students'GO	0.179	2.313	Accepted
Organization and clarity↔ students'GO	0.190	2.336	Accepted
Relevance ↔ students'GO	0.129	2.313	Accepted
students' perception on TP↔ students' GO	0.327	12.341	Accepted

The first hypothesis of the research states that "there is a causal association across students' perception on TP and GO. According to Table 7, it is observed that t-statistic related to the causal association across different elements of teaching practice and the total students' perception on TP is higher than 1.96, at a significance level of 0.05. Therefore, there is a causal association across components of TP (i.e. organization and clarity, feedback-feedforward and relevance) and GO. Furthermore, it is observed that the path coefficient and t-statistic pertaining to the causal association across students' perception on TP with GO are equal to 0.327 and 12.341, respectively, which shows that there is a causal association across students' perception on TP and GO.

The results of the analysis of casual association across students' perception on TP and learner achievements are summarized in Table 8 in detail.

Table 8

The Path Coefficient and T-Statistic for the Casual Association across Students' Perception on TP and Learner Achievements

H02	The path coefficient (β)	t Value	Result
Feedback-feed Forward ↔ students' vocabulary development	0.255	2.371	Accepted
Organization and clarity↔ students' vocabulary development	0.185	2.579	Accepted
Relevance ↔ students' vocabulary development	0.106	2.053	Accepted
students' perception on TP↔ students vocabulary development	0.408	15.514	Accepted

The Second hypothesis of the research states that "there is a causal association across students' perception on TP and learner vocabulary development. It was found that the value of t-statistic is higher than 1.96 for the interaction between components of students' perception on TP and students' vocabulary development, at a significance level of 0.05. This finding suggests that there is a causal association across these components and learners' affective and vocabulary development. According to the path coefficient and t-statistic pertaining to the causal association across students' perception on TP with learner achievements there is a causal association across students' perception on TP and learner achievements ($P<0.05$). It is noteworthy that organization and clarity was found to have higher interaction with students' vocabulary development. Moreover, relevance was found to have higher rate of interaction with students' affective achievements. Table 9 represents the results of the analysis of casual association across students' growth orientation and learner achievements.

Table 9

The Path Coefficient and T-Statistic for the Casual Association across Sstudents' Growth Orientation and Learner Achievements

H03	The path coefficient (β)	t Value	Result
Growth Mindset ↔ students' vocabulary development	0.274	10.254	Accepted
Self-based goals ↔ students' vocabulary development	0.265	7.568	Accepted
Task-based goals↔ students' vocabulary development	0.284	7.542	Accepted

Table 9 presents the path coefficient and t-statistic pertaining to the causal association between different elements of GO with learner vocabulary development. The results show that there is a causal association between GO elements and learners' affective and vocabulary development ($t \geq 1.96$, $p < 0.05$). Collectively, Table 9 indicates that students' vocabulary development has stronger association with task-based growth goals.

The third hypothesis related to the model tests the causal association across students' perception on TP and student achievements, considering the mediating role of GO. The Sobel test was used to test this hypothesis. The Sobel test statistic is as follows:

Table 10

Results of Sobel Test for the Mediating Role of Growth Orientation in the Causal Association Between Students' Perception on TP and Student Achievements

Variable	β	Z	VAF	Result
a (students' perception on TP → students' GO)	0.327			
b (students' GO → students' vocabulary development)	0.331			
S_a	0.026	8.933	0.210	Accepted
S_b	0.026			
c (students' perception on TP → students vocabulary development)	0.408			

According to Table 10 and the results of the Z statistic, it can be said with 99% confidence that the causal association across students' perception on TP and student achievements is mediated by GO (with a Z value of 8.933 and a VAF value of 0.210). Based on the results of the VAF (variance inclusion) test, it can be concluded that GO can explain and mediate 21% of the causal association between students' perception on TP and students' achievements.

Discussion

The present study aimed to situate Bandura's (1986, 1997) social cognitive theory of triadic reciprocal determinism in the Iranian EFL context. It examined how students' perceptions of TP, specifically instructional relevance, organization and clarity, and feedback/feedforward, interact with GO and possibly influence vocabulary development among secondary high school learners. Through partial least squares structural equation modeling on a large sample of 1,105 tenth-grade students, the results confirmed significant direct paths from TP to GO ($\beta = 0.327$, $t = 12.341$, $p < .001$) and from TP to vocabulary achievement ($\beta = 0.364$, $R^2 = 0.364$), with GO emerging as a significant mediator (Sobel $Z = 8.933$, $VAF = 0.210$). Notably, task-based growth goals within GO showed the strongest predictive power for vocabulary outcomes, while organization and clarity exhibited the most robust direct link to achievement. These findings not only validate the hypothesized model but also illuminate the dynamic interplay of environmental (i.e., TP), personal (i.e., GO), and behavioral (i.e., vocabulary development) factors, extending Bandura's framework by demonstrating its applicability in EFL setting.

Theoretically, the results refine SCT by specifying how instructional elements recalibrate personal factors in a culturally constrained environment. In Bandura's triadic model, reciprocal determinism posits that environmental influences shape personal cognitions, which in turn drive behavioral adaptation (Bandura, 1997). Here, feedback/feedforward emerges as the most potent environmental antecedent, fostering incremental beliefs and mastery-approach goals that buffer against fixed mindsets prevalent in exam-oriented systems (Dweck, 2006). This aligns with Elliot and McGregor's (2001) achievement goal framework, where task-based growth goals, emphasizing competence improvement against criteria, proved particularly adaptive for vocabulary acquisition, a domain requiring iterative practice and self-regulation. The mediation effect (21% variance accounted for) underscores GO's role as a transformative personal mechanism, challenging the unidirectional environmental determinism often critiqued in SCT applications to non-Western contexts (Schunk & DiBenedetto, 2020). By integrating growth constructs (mindset, self-based goals, task-based goals) into a unitary GO framework (Bostwick et al., 2017, 2019), the study advances Martin et al.'s (2022) growth goal-setting model, revealing how instructional clarity disrupts cycles of learned helplessness in EFL learners.

These patterns both replicate and extend prior research, while highlighting context-specific nuances. The direct positive associations between TP dimensions and vocabulary achievement corroborate meta-analytic evidence on instructional quality's role in language outcomes (Hulleman et al., 2017; Wisniewski et al., 2020), as well as Iranian studies showing clarity and relevance as predictors of EFL engagement (Khajavy et al., 2021). Similarly, GO's link to achievement echoes Bostwick et al. (2022), who found growth goals moderating adaptive engagement in mathematics, and Martin et al. (2020), who linked them to perseverance in secondary settings. However, the full mediation through GO diverges from direct-effect models in Western samples (e.g., Burns et al., 2019), where environmental influences often bypass personal factors. In the Iranian EFL context, this mediation may reflect cultural emphases on teacher authority and normative evaluation, where students' growth beliefs act as a necessary psychological filter for instructional benefits (Papi & Abdollahzadeh, 2012). Unlike studies in learner-centered systems (Granziera et al., 2022), where feedback effects are immediate, the amplified role of feedforward here aligns with feedback literacy models (Carless & Boud, 2018; Winstone et al., 2022), suggesting that Iranian learners, facing rote-heavy curricula, require explicit future-oriented guidance to translate feedback into growth-oriented action.

Conclusion

The findings hold profound relevance for EFL education system all around the world more particularly the Iranian EFL context, where secondary learners grapple with motivational and psychological challenges amid high-stakes national exams like the Konkoor, large class sizes, and a traditional emphasis on memorization over communicative competence (Rahimi & Weisi, 2018). In this context, where fixed mindsets and performance-avoidance goals contribute to widespread de-motivation and learned helplessness (Khajavy et al., 2018; Papi & Khajavy, 2021), the mediating power of GO offers a culturally attuned counter-narrative: even within resource-limited public schools, low-cost TP adjustments such as transparent lesson structuring and actionable feedforward can cultivate incremental beliefs, reducing dropout risks and enhancing vocabulary retention critical for exam success. This is especially vital for female-majority samples like the present one (66% female), as Iranian girls often face compounded gender-based motivational barriers in STEM-adjacent EFL tasks (Derakhshan et al., 2023). By positioning GO as a bridge, the study addresses a key gap in Iranian SCT applications, where prior work has focused on self-efficacy without integrating growth goals (Fathi et al., 2023).

Practically, these insights inform targeted interventions for Iranian EFL educators and policymakers. Teacher training under the Ministry of Education should emphasize modular workshops on feedforward techniques (e.g., personalized progress prompts during vocabulary drills) to foster task-based goals without overhauling curricula. Materials developers could embed reflective prompts and relevance-linking activities, promoting self-based growth in under-resourced rural-urban divides. At the policy level, incorporating GO metrics into teacher evaluations could incentivize clarity-focused practices, mitigating burnout and enhancing equity for underrepresented learners (Fathi et al., 2021). However, certain limitations temper these conclusions and point to future directions. The cross-sectional design limits causal inference; longitudinal studies tracking TP-GO-achievement trajectories over a semester are thus essential. In addition, self-reported TP perceptions may inflate shared method variance, warranting multi-source data (e.g., classroom observations). Finally, the focus on vocabulary in Northwest Iranian public schools constrains generalizability; replications in diverse provinces and outcomes (e.g., speaking proficiency) are needed.

All in all, this study advances Bandura's (1986) SCT by empirically situating triadic reciprocal determinism in EFL secondary classrooms more specifically in Iran, revealing GO as the key personal mediator through which teaching practices drive vocabulary achievement. Amid myriad challenges in educational system, the pronounced effects of feedback/feedforward and organizational clarity offer a pragmatic blueprint for reform. Fostering task-based growth goals can empower learners to reframe instruction as mastery opportunities, enhancing communicative equity. By bridging international theory with Iranian realities, these findings not only enrich global EFL scholarship but also equip educators, textbook designers, and policymakers with actionable strategies to cultivate resilient language learners. Future interventions testing integrated TP-GO training can ultimately nurture a generation of agentic, growth-oriented EFL achievers.

Acknowledgments

The authors would like to express their sincere gratitude to the participating schools, English teachers, and students whose cooperation and support made the data collection process possible. Their valuable contributions were essential to the successful completion of this study.

References

Bandura, A. (1986). *Social foundations of thought and action*. Prentice Hall.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes*, *50*(2), 248–287. https://doi.org/10.1016/0749-5978(91)90022-L

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W.H. Freeman.

Basso, D., & Olivetti Belardinelli, M. (2006). The role of the feedforward paradigm in cognitive psychology. *Cognitive Processing*, *7*(2), 73–88. https://doi.org/10.1007/s10339-006-0034-1

Blaich, C., Wise, K., Pascarella, E. T., & Roksa, J. (2016). Instructional clarity and organization: It's not new or fancy, but it matters. *Change: The Magazine of Higher Learning*, *48*(4), 6–13. https://doi.org/10.1080/00091383.2016.1198142

Brown, H. D. (2007). *Principles of language learning and teaching* (5th ed.). Pearson Education.

Bostwick, K. C. P., Collie, R. J., Martin, A. J., & Durksen, T. L. (2017). Students' growth mindsets, goals, and academic achievement in mathematics. *Zeitschrift Fur Psychologie / Journal of Psychology*, *225*(2), 107-116. https://doi.org/10.1027/2151-2604/a000287

Bostwick, K. C. P., Martin, A. J., Collie, R. J., & Durksen, T. L. (2019). Growth orientation predicts gains in middle and high school students' mathematics achievements over time. *Contemporary Educational Psychology*, *58*, 213–227. https://doi.org/10.1016/j.cedpsych.2019.03.010

Bostwick, K. C. P., Collie, R. J., Martin, A. J., & Durksen, T. L. (2020). Teacher, classroom, and student growth orientation in mathematics: A multilevel examination of growth goals, growth mindset, engagement, and achievement. *Teaching and Teacher Education*, *94*, 103100. https://doi.org/10.1016/j.tate.2020.103100

Bostwick, K. C. P., Collie, R. J., Martin, A. J., & Durksen, T. L. (2022). Motivation in Context: A Multilevel Examination of Growth Orientation across One Year. *Journal of Applied Developmental Psychology, 81,* 101435. https://doi.org/10.1016/j.appdev.2022.101435

Burns, E. C., & Martin, A. J. (2018). Synergistic relationships between growth mindset, academic buoyancy, and academic achievement in mathematics. *International Journal of Educational Research, 92*, 1–10. https://doi.org/10.1016/j.ijer.2018.08.005

Burns, E. C., Martin, A. J., & Collie, R. J. (2018). Adaptability, personal best (PB) goal setting, and gains in students' academic achievement: A longitudinal examination from a social cognitive perspective. *Contemporary Educational Psychology*, *53*, 57–72. https://doi.org/10.1016/j.cedpsych.2018.02.001

Burns, E. C., Martin, A. J., & Evans, P. A. (2019). The role of teacher feedback–feedforward and personal best goal setting in students' mathematics achievement: a goal setting theory perspective. *Educational Psychology*, *41*(7), 1–19. https://doi.org/10.1080/01443410.2019.1662889

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Derakhshan, A., Kruk, M., Mehran, P., & Tahernia, M. (2023). An investigation of the Iranian EFL learners' FL enjoyment, anxiety, and attitude/motivation in online and traditional classes. *Computer Assisted Language Learning, 36*(7), 1325–1351. https://doi.org/10.1080/09588221.2021.1983045

Duran, D., & Kääntä, L. (2023). Building word knowledge through integrated vocabulary explanations in ESL tutorials. *Linguistics and Education, 75*, Article101182. https://doi.org/10.1016/j.linged.2023.101182

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development*. Psychology Press.

Dweck, C. S. (2006). *Mindset: The new psychology of success.* Random House.

Elliot, A. J., & McGregor, H. A. (2001). A 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology, 80*(3), 501–519. https://doi.org/10.1037/0022-3514.80.3.501

Fathi, J., Greenier, V., & Derakhshan, A. (2021). Teacher self-efficacy, reflection, and burnout among Iranian EFL teachers: The mediating role of emotion regulation. *Iranian Journal of Language Teaching Research, 9*(2), 85–104. https://doi.org/10.30466/ijltr.2021.121346

Fathi, J., Greenier, V., & Derakhshan, A. (2023). Teacher efficacy, reflection, and burnout among Iranian EFL teachers: The mediating role of emotion regulation. *Iranian Journal of Language Teaching Research, 11*(2), 85–104. https://doi.org/10.30466/ijltr.2023.121346

Granziera, H., Liem, G. A. D., Chong, W. H., Martin, A. J., Collie, R. J., Bishop, M., & Tynan, L. (2022). The role of teachers' instrumental and emotional support in students' academic buoyancy, engagement, and academic skills: A study of high school and elementary school students in different national contexts. *Learning and Instruction*, *80*, Article 101619. https://doi.org/10.1016/j.learninstruc.2022.101619

Hajabi, M., Taheri, S., Fahandezh, F., & Salari, H. (2018). The role of web-based language teaching on vocabulary retention of adult pre-intermediate EFL learners. *Journal of Language Teaching and Research*, *9*(2), 372. https://doi.org/10.17507/jltr.0902.20

Hulleman, C. S., Kosovich, J. J., Barron, K. E., & Daniel, D. B. (2017). Making connections: Replicating and extending the utility value intervention in the classroom. *Journal of Educational Psychology, 109*(3), 387–404. https://doi.org/10.1037/edu0000146

Khajavy, G. H., Bardach, L., Hamedi, S. M., & Lüftenegger, M. (2018). Broadening the nomological network of classroom goal structures using doubly latent multilevel modeling. *Contemporary Educational Psychology*, *52*, 61-73. https://doi.org/10.1016/j.cedpsych.2017.10.004

Khajavy, G. H., MacIntyre, P. D., & Hariri, J. (2021). A closer look at grit and language mindset as predictors of foreign language achievement. *Studies in Second Language Acquisition, 43*(2), 379–402. https://doi.org/10.1017/S0272263120000480

Lei, H., Cui, Y., & Chiu, M. M. (2017). The relationship between teacher support and students' academic emotions: A meta-analysis. *Frontiers in Psychology*, *8*, Article 2288. https://doi.org/10.3389/fpsyg.2017.02288

Martin, A. J. (2006). Personal bests (PBs): A proposed multidimensional model and empirical analysis. *British Journal of Educational Psychology*, *76*, 803–825. http://dx.doi.org/10.1348/000709905X55389

Martin, A. J. (2015). Growth approaches to academic development: Research into academic trajectories and growth assessment, goals, and mindsets. *British Journal of Educational Psychology*, *85*, 133–137. https://doi.org/10.1111/bjep.12071

Martin, A. J., Burns, E. C., Collie, R. J., Bostwick, K. C. P., Flesken, A., & McCarthy, I. (2022). Growth goal setting in high school: A large-scale study of perceived instructional support, personal background attributes, and engagement outcomes. *Journal of Educational Psychology, 114*(4), 752–771. https://doi.org/10.1037/edu0000682

Martin, A. J., & Evans, P. (2018). Load reduction instruction: Exploring a framework that assesses explicit instruction through to independent learning. *Teaching and Teacher Education*, *73*, 203–214. https://doi.org/10.1016/j.tate.2018.03.018

Martin A. J, & Evans P. (2019). Load reduction instruction (LRI): Sequencing explicit instruction and guided discovery to enhance students' motivation, engagement, learning, and achievement', In S. Tindall-Ford, S. Agostinho, & J. Sweller. *Advances in cognitive load theory: Rethinking teaching* (pp. 15 – 29). Routledge.

Martin, A. J., Ginns, P., & Collie, R. J. (2023). University students in COVID-19 lockdown: The role of adaptability and fluid reasoning in supporting their academic motivation and engagement. *Learning and Instruction*, *83*, Article 101712. http://dx.doi.org/10.1016/j.learninstruc.2022.101712

Martin, A. J., Ginns, P., Burns, E. C., Kennett, R., & Pearson, J. (2021). Load reduction instruction in science and students' science engagement and science achievement. *Journal of Educational Psychology, 113*(6), 1126–1142. https://doi.org/10.1037/edu0000552

Nation, I. S. P., & Newton, J. (1997). Teaching vocabulary. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in Language Teaching* (pp. 100-114). Cambridge University Press.

Nation, P. (2001). *Learning vocabulary in another language*. Cambridge University Press. https://dx.doi.org/10.1017/CBO9781139524759

Papi, M., & Abdollahzadeh, E. (2012). Teacher motivational practice, student motivation, and possible L2 selves: An examination in the Iranian EFL context. *Language Learning, 62*(2), 571–594. https://doi.org/10.1111/j.1467-9922.2011.00632.x

Papi, M., & Khajavy, G. H. (2021). Motivational mechanisms underlying second language achievement: A regulatory focus perspective. *Language Learning, 71*(2), 537–572. https://doi.org/10.1111/lang.12443

Rahimi, M., & Weisi, H. (2018). Reflective practice, self-efficacy and research practice of EFL teachers: Examining possible relationships. *Issues in Educational Research, 2018, 28*(3), 756-780.

Schmidt, J. A., Shumow, L., & Kackar-Cam, H. (2015). Exploring teacher effects for mindset intervention outcomes in seventh-grade science classes. *Middle Grades Research Journal*. *10*(2), 17–32. https://doi.org/10.1108/MGRJ-10-2015-0004

Schunk, D. H., & DiBenedetto, M. K. (2020). *Motivation and social cognitive theory*. Motivation and Social Cognitive Theory. *Contemporary Educational Psychology, 60*, Article 101832. https://doi.org/10.1016/j.cedpsych.2019.101832

Tenenhaus, M., Amato, S., & Vinzi, V. E. (2004) A global goodness-of-fit index for PLS structural equation modelling. *Proceedings of the XLII SIS Scientific Meeting, 1*, 739-742.

Winstone, N. E., Ajjawi, R., Dirkx, P., & Boud, D. (2022). *Feedback literacy: Academic institutions need curriculum reform*. In C. Carless & D. Boud (Eds.), *The Power of Assessment for Learning* (pp. 1–15). Routledge.

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology, 10*, Article 3087. https://doi.org/10.3389/fpsyg.2019.03087

Zeng, G., Chen, X., Cheung, H. Y., & Peng, K. (2019). Teachers' growth mindset and work engagement in the Chinese educational context: Well-being and perseverance of effort as mediators. *Frontiers in Psychology*, *10*. Article 839. https://doi.org/10.3389/fpsyg.2019.00839

Appendix A

Growth Orientation and Teaching Practice Inventory

دانش آموز عزیز، سلام

خواهشمند است با اختصاص دقایقی از وقت خود، ما را در انجام این پژوهش یاری نمایید لازم به ذکر است که پاسخ های شما بصورت محرمانه و بی نام ثبت خواهد شد و نتایج آن تنها برای اهداف پژوهشی استفاده می گردد. پیشاپیش از همکاری شما صمیمانه تشکر می نماییم.

نام دبیر:	نام مدرسه:	جنسیت:
نام کلاس:	شهر:	سن:

جهت پاسخگویی لطفا دیدگاه خود را نسبت به درس زبان انگلیسی، کلاس درسی مدرسه و دبیر مربوطه در نظر گرفته و از راهنمای زیر کمک بگیرید:

کاملا مخالف:۱ مخالف: ۲ نظری ندارم: ۳ موافق:۴ کاملا موافق: ۵

ردیف	عبارات	۱	۲	۳	۴	۵
۱	وقتی مطلب جدیدی در کلاس آموزش داده می شود، معلم اطمینان حاصل می کند که در ابتدا با قسمت های آسان تر مبحث را شروع کند.					
۲	در حین کار روی تمرین ها و فعالیت ها در کلاس، معلم به خوبی به ما کمک می کند.					
۳	در کلاس انگلیسی معلم اطمینان حاصل می کند که قبل از وارد شدن به مبحث جدید به اندازه کافی مبحث قبلی را تمرین کرده ایم.					
۴	در کلاس انگلیسی معلم بازخورد های سازنده ای ارائه می دهد که در بهبود یادگیری و انجام فعالیت ها کمک می کند.					
۵	به محض اینکه متوجه مبحث شدیم معلم به ما این فرصت را می دهد که بطور مستقل روی فعالیت ها کار کنیم.					
۶	معلم انگلیسی توضیحات واضح ارائه می کند.					
۷	معلم ما از مثال ها و تصویر سازی ها برای توضیح دادن نکات سخت به خوبی استفاده می کند.					
۸	معلم ما به طور موثری مطالب را مرور و خلاصه می کند.					
۹	معلم ما بطور واضح مفاهیم انتزاعی و نظریه ها را تفسیر می کند.					
۱۰	معلم ما تکلیف هایی می دهد که به یادگیری ماده درسی کمک می کند.					
۱۱	ارائه مطالب توسط معلم کامل سازمان یافته است.					
۱۲	معلم در کلاس کاملا آماده است.					
۱۳	معلم ما از مدت زمان کلاسی بطور کارآمد استفاده می کند.					
۱۴	اهداف درس و پیش نیازهای آن کاملا واضح توضیح داده می شوند.					
۱۵	معلم روی مطالبی که تدریس می کند تسلط کامل دارد.					

					محتوای درسی مفید است.	۱۶
					محتوای درسی معنادار است.	۱۷
					محتوای درسی هدفمند است.	۱۸
					میزان هوش شما ثابت است و شما کار چندانی برای تغییر آن نمیتوانید انجام دهید	۱۹
					شما همواره به شکل چشم گیری قادر به تغییر میزان هوش خود هستید.	۲۰
					من هدف های مشخصی برای تکالیف مدرسه تعیین می کنم.	۲۱
					من هدف های چالش برانگیزی برای انجام تکالیف مدرسه برای خودم تعیین می کنم.	۲۲
					من بیش از اینکه با سایر دانش آموزان در رقابت باشم، با عملکردهای قبلی خودم رقابت می کنم.	۲۳
					سعی می کنم تکالیف مدرسه ام را بهتر از قبل انجام دهم.	۲۴
					هدف من در کلاس گرفتن نمره های بهتر نسبت به سایر دانش آموزان است.	۲۵
					گاهی اوقات از این می ترسم که شاید محتوی این کلاس را با آن دقتی که می خواهم متوجه نشوم.	۲۶
					من می خواهم به طور کامل بر مطالب ارائه شده در این کلاس تسلط داشته باشم.	۲۷
					ترس من از عملکر ضعیف در این کلاس اغلب به من انگیزه می دهد.	۲۸

Appendix B

Vocabulary Test

1 I can't understand why some people .... or kill the animals.
a. protect b. save c. hurt d. lose
2 I think that shopkeepers will soon .... the price from $50 to $60.
a. cut b. increase c. pay d. divide
3 I love Swiss chocolate, ... dark Swiss chocolate.
a. naturally b. hopefully c. especially d. orally
4 No one knows if they are alive or dead. We can only .... and wait.
a. choose b. enjoy c. hope d. explain
5 These animals sleep in the day and .... at night.
a. lose b. hope c. add d. hunt
6 Thanks for helping me with my homework; you .... my life!
a. destroyed b. hurt c. finished d. saved
7 I read on a website that dinosaurs .... about 65 million years ago. Is that true?
a. died out b. lost c. put out d. killed
8 Of course I make mistakes, I'm only ....
a. hunter b. plain c. human d. future
9 When they started their life together, they were young and full of ....
a. hope b. past c. family d. relatives
10 Firefighters came to .... the fire in the city center.

a. pay attention b. take care c. cut down d. put out

11 There are lots of things in that shop, for ...., games, toys and flowers.

a. different b. future c. example d. recently

12 I'm free tonight. I'm going to read some ....

a. dangers b. poems c. types d. plains

13 Many houses, schools and offices were damaged and thousands of .... people were taken to the hospital.

a. dangerous b. hopeful c. endangered d. injured

14 I really like to be a .... because I enjoy visiting different places in the world.

a. tourist b. hunter c. pilot d. zookeeper

15 The children were saved but the fire .... their home.

a. caught b. suffered c. destroyed d. injured

16 I don't feel like going to cinema. Let's take a walk ....

a. but b. instead c. rather d. short

17 They found the children .... and well after being missing for several days.

a. brave b. regular c. alive d. injured

18 Every time an animal ...., we lose a part of our natural world forever.

a. puts out b. grows up c. dies out d. takes off

19 I think it's .... for a mother to feel sad when her children leave home.

a. endangered b. natural c. wonderful d. amusing

20 It's important to ....care of your teeth.

a. grow b. make c. give d. take

21 The big panda is one of the world's most .... animals.

a. hopeful b. divided c. extra d. endangered

22 Tom .... his left knee during practice, so John had to play the game in his place.

a. hurt b. took c. gave d. left

23 She planned to work till she .... the money to buy a new laptop.

a. reported b. ended c. saved d. protected

24 There are a lot of dark .... in the sky. It's going to rain.

a. images b. sites c. clouds d. tickets

25 It is important to .... your skin from the harmful effects of sun.

a. attract b. protect c. divide d. destroy

26 She'll be happy to .... that she can leave hospital tomorrow.

a. hear b. describe c. lose d. increase

27 Like many young men, he does not .... attention to his health.

a. do b. make c. get d. pay

28 We have a/ an .... right now. Would you please wait here?

a. danger b. pattern c. average d. visitor

29 Everyone says Kate is very hard-working and will have a successful ....

a. idea b. past c. hope d. future

30 He isn't working today, so .... he'll feel better tomorrow.

a. carefully b. recently c. interestingly d. hopefully

Authors' Biography

Mohaddeseh Hashemi-Aghdam is a PhD candidate in TEFL/TESOL at the University of Urmia. She received her Bachelor's degree from teacher training college of Zanjan and her Master's degree in TEFL from the University of Tabriz and has over a decade of experience teaching English in various language institutes. Currently, she serves as an instructor at the Tabriz University of Medical Sciences and Payam-Noor University. Her primary research interests lie in second language acquisition, and language pedagogy. She is particularly interested in psycholinguistics.

Sima Modirkhamene has a PhD in TEFL/TESOL from the University of Surrey, UK. Since her return to Iran in September 2006, she has been lecturing at undergraduate and postgraduate levels and researching in Urmia University, Iran. Her main research interests include bilingualism, first and second language acquisition, cross-linguistic transfer, and multiple intelligences in relation to language learning. She has several paper and book publications in top-ranked publishers. She has served as the Head of the Department of English Language and Literature, college administrator, and the Deputy Head of the Foreign Languages Center of Urmia University for several years.

Parviz Alavinia holds a PhD in TEFL/TESOL from Allameh Tabataba'i University in Tehran. He has gotten his BA degree from Urmia University and his MA from Kharazmi University in Tehran. He has been instructing in several language schools for about 22 years, and has lectured at Urmia university since about 15 years ago. Successive to the completion and fulfillment of his PhD degree (February, 9, 2010), he has been involved as a full-time associate professor and staff member at Urmia University. His main areas of interests include: psycholinguistics, philosophy of language, teacher education, emotional intelligence, motivation and TBLT.

Volume 14, Issue 1
Winter and Spring, 2026
pp. 247-264

Journal of Applied Linguistics and Applied Literature: Dynamics and Advances

Resilience amid Ruin: A Postcolonial Ecocritical Study of Ngũgĩ wa Thiong'o's Petals of Blood

Reyhaneh Goudarzi[1] and Hossein Nazari[2,*]

[1]*MA in English Literature, University of Tehran, Tehran, Iran*
Email: goudarzi.reyhaneh@ut.ac.ir

[2]*Corresponding Author: Assistant Professor of English Language and Literature, Department of English Language and Literature, Faculty of Foreign Languages and Literatures, University of Tehran, Tehran, Iran*
ORCID: 0000-0001-6429-3779; Email: nazarih@ut.ac.ir

Abstract

This study offers a postcolonial ecocritical critique of Ngũgĩ wa Thiong'o's Petals of Blood (1977) within the context of postcolonial Kenya's socio-economic crises, exacerbated by colonial legacies and ongoing neocolonial influences. The analysis examines the novel's portrayal of resilience and trauma amid environmental and socio-economic turmoil. We argue that the novel presents a microcosm of Kenya's socio-political landscape, highlighting the trauma inflicted by (neo-)colonial exploitation. The central characters embody this collective trauma as they confront oppression, violence, and betrayal. The environmental deterioration of Ilmorog serves as an example of colonial agricultural practices and the mismanagement pervasive in the postcolonial era. This study argues that indigenous knowledge and ecological practices offer pathways to healing and resilience, contrasting sharply with profit-driven approaches. Nature symbolizes hope and resistance, while the blood motif represents violence and communal unity. Ultimately, the novel reflects the interplay of destruction and regeneration in postcolonial Kenya, advocating for the interconnectedness of ecological and social justice. It emphasizes the importance of reclaiming indigenous knowledge and sustainable practices for reconciliation and resilience.

Keywords: ecocriticism, resistance, environmental degradation, Petals of Blood, postcolonial Kenya

ARTICLE INFO
Research Article
Received: Friday, December 26, 2025
Accepted: Wedensday, March 18, 2026
Published: Wedensday, April 1, 2026
Available Online: Wedensday, March 18, 2026
DOI: https://doi.org/ 10.22049/jalda.2026.31347.1944

Online ISSN: 2821-0204; Print ISSN: 28208986

Introduction

This study embarks on an in-depth critique of *Petals of Blood* (1977), which vividly portrays the sociopolitical upheavals in postcolonial Kenya and reflects the ecological degradation that parallels the nation's quest for identity and justice. Ngũgĩ employs environmental metaphors to offer a unique perspective on the interplay between human resilience and ecological trauma.

Set against a newly independent Kenya grappling with the remnants of colonialism, the novel portrays the challenges of constructing a new national identity. Ngũgĩ uses the environment as a narrative device to underscore the Kenyans' broader sociopolitical and economic struggles. By portraying the land and its degradation, he critiques the neocolonial exploitation plaguing the nation, illustrating how environmental trauma is intertwined with the human condition in postcolonial contexts.

The transformation of Ilmorog, a rural village, into an industrialized town symbolises Kenya's broader environmental and social transformations. Employing a postcolonial ecocritical framework, this study critiques how Ngũgĩ intertwines the themes of environmental degradation and human resilience to reflect the complexities of postcolonial Keny. Central to this analysis is ecological and human resilience, providing insights into how communities navigate the lingering impacts of colonial exploitation and neocolonial practices.

Through a close reading of the novel, the study investigates how Ngũgĩ's narrative strategies highlight the interconnectedness of ecological and social trauma. The novel presents a rich tapestry of characters who grapple with personal and collective traumas intimately connected to the land and its exploitation. Their stories are intertwined with the history of post/colonial Kenya, reflecting enduring legacies of displacement, economic disenfranchisement, and cultural erasure. Our critique explores how the text's portrayal of ecological degradation illustrates both colonial and postcolonial exploitation and how the vision of resilience offers a pathway toward healing and regeneration. Additionally, it highlights how literature can effectively articulate the complex interplay between human and ecological systems and envision more just and sustainable futures in postcolonial contexts.

Environmental Degradation

In *Petals of Blood*, Ngũgĩ uses environmental degradation as a metaphor for neocolonialism's social, political, and psychological impacts in post-independence Kenya. Ilmorog's transformation from a vibrant rural community to a barren wasteland symbolizes the exploitation and betrayal of the Kenyan people.

Ngũgĩ mirrors Kenya's economic exploitation through Ilmorog's physical destruction. Industrialization and capitalist ventures displace villagers, stripping them of autonomy and resources. The shift from a self-sustaining community to a landscape dominated by external forces becomes a metaphor for how elites and foreign interests co-opt Kenya's resources. This is evident in the villagers' alienation from their land: "The land... had once belonged to the people, now it was owned by a few, fenced and marked... they are our new masters, and we are their slaves" (236). The enclosure of communal land symbolizes broader economic subjugation, while the villagers themselves become "slaves" under capitalism, reproducing colonial exploitation.

The devastating drought in Ilmorog is more than an environmental calamity—it mirrors the villagers' cultural and spiritual loss. Traditional customs disintegrate as their connection to the land weakens: "Without rain, the land cracked, and so did the spirits of the people. The old stories and songs had no meaning in this wasteland" (p. 132). The "cracked" land represents the fragmentation of the community's spirit, and the loss of stories reflects the erosion of cultural heritage as capitalist forces disrupt their relationship with the land.

Ngũgĩ traces Ilmorog's decline into a wasteland as a metaphor for Kenyan society's decay under neo-colonialism. The transformation reflects the unfulfilled promises of independence, where local elites replicate colonial exploitation. The contrast between Ilmorog's past and present underscores the degradation of Kenya's moral and social fabric: "Ilmorog had been fertile, with green hills and clear waters... now, it was all barren, dust-covered, and abandoned by life" (p. 65). The shift metaphorically represents how post-independence potential has withered under corruption and greed.

As Ilmorog's environment deteriorates, so does the villagers' psychological state. Ngũgĩ connects the land's ruin with the community's emotional suffering. The environmental collapse symbolises the villagers' powerlessness and despair: "The sun was scorching, no crops could grow, and the people despaired. Some even said the gods had forsaken them" (p. 128). The "scorching" sun and barren land reflect physical hardship and psychological trauma, as the villagers feel forsaken and powerless in the face of larger systemic forces.

The Destruction of Land: Neocolonial Betrayal

Throughout the novel, the destruction of the land is used as a metaphor for the betrayal of Kenya's independence movement. While independence promised freedom from colonial exploitation, the new Kenyan elite instead perpetuated the same systems of oppression. This betrayal is reflected in the degradation of Ilmorog's environment, as local elites and foreign companies exploit the land for their gain. The transformation of Ilmorog into an industrial town, exemplify the betrayal of the people's hopes for a better future: "Now cement and steel were threatening to crush the very soul of the place" (278). The imagery of "cement and steel" crushing the "soul" of Ilmorog underscores how capitalist exploitation destroys not just the physical landscape but also the cultural and spiritual essence of the village.

The stark image of a once-fertile land turning barren also represents the broader consequences of neocolonial exploitation. As the land is depleted by plantations and industries, so are the people exploited for their labor. The environmental degradation serves as a reminder that both the land and its people are being drained of their vitality and potential by forces of capital: "The fields once teemed with life, now all that remained was dead soil and poisoned water" (p. 233), exemplifying how the resources of Kenya are being sucked dry by exploitative practices, mirroring the exploitation of the people. This environmental ruin metaphorically represents how neocolonial structures continue to oppress and deplete both the land and the population, perpetuating a cycle of degradation and trauma.

Psychological Impact on Characters

The environmental trauma in Ilmorog profoundly affects the mental health of characters. In *Petals of Blood*, Ngũgĩ presents four central figures, whose lives converge in the village: Munira, an idealistic teacher; Karega, a passionate political activist; Wanja, a determined yet vulnerable woman; and Abdulla, a disillusioned Mau Mau veteran.

Seeking redemption, Munira arrives in Ilmorog hoping to educate and uplift the impoverished village. However, his idealism crumbles as he witnesses the environmental decay and villagers' indifference. The harsh realities of poverty, neglect, and capitalist forces shatter his belief in education's power to transform lives. This disillusionment, coupled with his personal guilt and unresolved moral conflicts, leads to his descent into madness, reflecting how environmental and social degradation affect his psyche.

Unlike Munira, Karega is driven by a desire for social justice and is acutely aware of the systemic exploitation of the villagers. His political consciousness stems from personal losses, such as the death of his brother, which fuel his commitment to fighting the neocolonial and capitalist forces destroying Ilmorog. However, this awareness also brings psychological distress, as he grapples with the overwhelming challenge of mobilizing a fragmented community. His activism reflects a broader response to the trauma of environmental and social injustice, as he seeks to restore balance by fighting for workers' rights and environmental justice.

Wanja returns to Ilmorog after enduring personal and economic hardships in the city, including betrayal and sexual exploitation. The degradation of the land mirrors her psychological turmoil, as she struggles with a fractured sense of identity. Her choice to run a bar and later engage in prostitution highlights the limited options available to her in a society that commodifies both women and land. Despite these challenges, Wanja's resilience and determination to rebuild her life suggest that reconnecting with the land offers psychological recovery and empowerment potential.

Abdulla, veteran disabled from injuries sustained in the fight for independence, represents the disillusionment of the post-colonial era. He runs a small shop in Ilmorog but feels betrayed by the neocolonial order, which has failed to deliver on the promises of freedom and equality. The environmental devastation of Ilmorog exacerbates his psychological distress, as the erosion of communal values and land mirrors his sense of loss and futility.

Together, these characters' experiences illustrate the profound impact of environmental and social degradation on mental health. Ilmorog's physical decline acts as a catalyst for unraveling both individual and communal well-being, emphasizing the deep connection between the land's health and the emotional stability of its inhabitants.

Socio-Economic Consequences

In *Petals of Blood*, the socio-economic effects of environmental degradation are dramatic, intensifying the trauma experienced by the community. The transformation of Ilmorog into a commercial hub driven by neocolonial interests results in resource depletion, rendering villagers economically precarious and socially divided.

As Berkes and Berkes argue, modernization and industrialization undermine traditional ecological knowledge, destabilizing indigenous communities (p. 6). Throughout the novel, this is evident in how commercial encroachment results in unemployment, poverty, and alienation. Ilmorog, once a remote agricultural village thriving on land-based practices, faces disruption as external forces invade. A pivotal moment occurs with the initiation of the Trans-Africa Highway, initially regarded as a sign of progress but ultimately precipitating Ilmorog's environmental and social decline. Although the highway aims to connect Ilmorog to economic centers, it instead serves as a vehicle for exploitation by business elites and corrupt politicians. Villagers are coerced into selling their ancestral lands to developers, resulting in considerable unemployment and the abandonment of traditional farming. The land is transformed into commercial space, depriving the villagers of both their means of survival and their cultural heritage.

The shift to a commercial economy further marginalizes the villagers. New businesses like the brewery introduce an economy where locals struggle to compete, face poverty and cannot find stable work. This economic shift creates a cultural and social rift, undermining the communal bonds that once unified Ilmorog.

A poignant example of commercialization's destructive impact is Nyakinyua, Wanja's grandmother. After losing her land due to new taxes, Nyakinyua's land, is consumed by commercial interests, leading to her death. Villagers like Wanja, who turns to running a bar and brothel for survival, highlight the socio-economic trauma caused by environmental degradation and the loss of ecological knowledge.

The novel represents collective trauma, portraying how environmental degradation creates a shared injury for the community. The villagers' journey to Nairobi represents a united effort to find justice. They want the government to acknowledge their struggles with the worsening drought, failing crops, and rising commercial pressures. While the pilgrimage highlights their resilience, their arrival in Nairobi reveals indifference and hostility, amplifying their feelings of betrayal and alienation. This symbolic and physical journey exemplifies the villagers' courage in facing systemic neglect.

As Ilmorog villagers embark on their pilgrimage in a desperate plea for government help, they could consider other resistance and survival strategies. When migrating to urban areas, they often face increased marginalization, bureaucratic indifference, and hostility instead of the anticipated support. Their plight highlights the systematic neglect faced by rural communities in a neocolonial state. However, in addition to migration, various forms of collective action and self-sufficiency could offer alternative routes to resilience and empowerment.

One possible strategy involves grassroots mobilization and self-organized resistance within Ilmorog. Rather than depending on the government, villagers could create a cooperative agricultural system, strengthening traditional farming methods that support the community while countering the invasion of capitalist enterprises. This strategy resonates with Vandana Shiva's emphasis on the value of indigenous agricultural knowledge in combating corporate exploitation and ecological harm (42). By reclaiming control over their land and resources, villagers could lessen their reliance on external entities that have traditionally disappointed them.

Moreover, forming partnerships with other marginalized rural communities could bolster their efforts. If the villagers engaged with nearby areas facing similar exploitation, they could create a more extensive and unified resistance movement. This larger coalition might pressure authorities and landowners through strikes, land occupations, or boycotts, akin to the peasant movements observed in postcolonial Kenya. As Ngũgĩ points out in his *Decolonising the Mind* (1986), genuine liberation emerges not from solitary migration but from collective action against oppressive systems (p. 88).

An alternative to migration involves education and raising political awareness within Ilmorog. Karega's role in the novel highlights the importance of political education in building resistance. Instead of abandoning the village, he could facilitate literacy and political workshops, equipping villagers with knowledge about their rights and the exploitation they endure. As Fanon asserts, "political education is an absolute necessity for colonized people; it is a prerequisite for developing a new consciousness" (p. 147). By understanding neocolonial structures, villagers could formulate strategies to reclaim their economic and political power without leaving their homeland.

Finally, it is also important to pursue legal and political engagement. Although the novel depicts the postcolonial government as corrupt and indifferent, the villagers can still organize petitions, hold peaceful protests, or seek legal action regarding their land rights. While this approach does involve risks, it provides a way to challenge the status quo within the current system.

Consequently, although migration seems essential for the villagers of Ilmorog, options like grassroots mobilization, coalition-building, education, and legal action might offer more sustainable paths for resistance. Ngũgĩ's depiction of Ilmorog's struggles emphasizes the necessity for systemic change, and these alternatives illustrate how resilience can be cultivated within the community rather than through forced relocation.

As G. D. Killam notes, "[t]he communal efforts in *Petals of Blood* to address environmental and social injustices highlight the significance of solidarity and collective action in healing from trauma" (p. 115). The villagers' unity in adversity illustrates the power of collective healing. Ngũgĩ emphasizes the interconnectedness of environmental and social injustices, advocating for solidarity, resilience, and activism as essential to addressing and recovering from trauma.

Intergenerational Trauma and Environmental Injustice

Petals of Blood illustrates environmental degradation as trauma the land and its people have suffered. The devastation of Ilmorog represents the erosion of ecological health and cultural identity caused by colonial exploitation and postcolonial mismanagement. The characters' profound bond with the land intensifies their sorrow over its deterioration, reflecting their trauma linked to historical injustices. A central theme in the novel is intergenerational trauma— the passing down of historical oppression through generations. Ngũgĩ conveys how the remnants of colonialism continue to haunt the descendants of those who directly endured its violence. The central characters personify the psychological and emotional wounds inherited from prior generations.

Abdulla, the Mau Mau veteran, embodies the lingering trauma of colonization. His injured leg serves not only as a physical reminder of history but also as a representation of ongoing pain and unresolved conflict. Catherine A. Odora Hoppers observes that "[t]he trauma of colonisation is not confined to the immediate victims but extends to subsequent generations, who inherit the psychological scars and socio-economic disadvantages" (p. 56). Abdulla's resentment and disillusionment are manifestations of this inherited trauma as he confronts the betrayal of the independence movement's promises.

Similarly, Wanja's life is marked by the repercussions of colonial exploitation. Her journey from a hopeful schoolgirl to a disillusioned barmaid and eventually a brothel madam illustrates the cyclical nature of trauma and exploitation. Ngũgĩ uses the character Wanja to highlight how "the wounds of the past manifest in the present, affecting the community's ability to heal and progress" (Mugo, 1978, p. 112).

Another consequence of colonial rule is the erosion of traditional culture and values, which exacerbates intergenerational trauma. Characters like Nyakinyua, the village elder, symbolize the cultural wisdom that is being lost. Her despair over the changing landscape speaks to the psychological impact of cultural disintegration. As the novel depicts, the disconnect experienced by the younger generation from their heritage further alienates them, intensifying their psychological distress.

The psychological impact of this cultural erosion is significant. The characters' struggles with the questions of identity and belonging are symptomatic of a more profound existential crisis precipitated by the (neo-)colonial disruption of their social fabric. This aligns well with Fanon's assertion that "colonialism is not satisfied merely with holding a people in its grip and emptying the native's brain of all form and content. By a kind of perverted logic, it turns to the past of the oppressed people and distorts, disfigures, and destroys it" (p. 210).

The novel vividly portrays the environmental degradation driven by the interests of multinational corporations and corrupt local elites, who prioritize profit over the well-being of the environment and the people. This is exemplified by the arrival of foreign investors and their local collaborators, who are merely concerned with extracting wealth from the land. The introduction of commercial farms, which focus on cash crops for export rather than subsistence agriculture, leads to the depletion of soil fertility and the disruption of local food systems. For instance, the brewery established in Ilmorog is owned by powerful business interests with ties to international markets. This business not only "diverts precious water resources away from the villagers' traditional farming activities, but also contributes to the pollution of the environment" (Ngũgĩ, p. 295).

The highway project is a large-scale infrastructure initiative to construct a major road through Ilmorog. Officially touted as a catalyst for regional development, the highway is positioned as a means to enhance transportation and economic opportunities. However, its true intention is to facilitate resource extraction by multinational corporations.

Kimeria, a prosperous and influential businessman, personifies the corrupt local elite who gains from this project. Utilizing his influence, Kimeria orchestrates deals selling communal land to these corporations. His actions epitomize the broader neocolonial exploitation depicted in the novel. While the highway project is marketed as progress, it ultimately results in environmental degradation, displacement of the local population, and converting their land into commercial estates. Consequently, the villagers are left with depleted land, underscoring how local and global forces collaborate to perpetuate exploitation and marginalization (p. 215).

The novel delves into the detrimental impact of neocolonial economic policies, portraying how profit-driven motives contribute to the systematic devastation of the environment and the traditional means of living. It depicts a narrative where multinational corporations prioritize profit at the expense of environmental sustainability and the well-being of the people. This results in a vicious cycle of environmental degradation and socio-economic decline, leaving the once-thriving land uninhabitable and perpetuating the historical legacy of colonial exploitation through the lens of contemporary global capitalism and its complicit local actors.

The construction of the Trans-Africa Road is a pivotal event in the novel that symbolises such exploitation. Initially seen as a symbol of progress, the road instead brings destruction, disrupting traditional ways of life and leading to environmental decay. The droughts that plague Ilmorog directly result from the ecological imbalance caused by deforestation and industrial activities. This scenario aligns with Shiva's observation that "development projects often result in the systematic destruction of ecosystems, leading to long-term ecological crises and the displacement of indigenous communities" (p. 42).

The environmental changes in Ilmorog have profound social implications, exacerbating existing inequalities and creating new forms of trauma and suffering. The loss of arable land and water sources forces many villagers to migrate to urban areas for work, only to encounter further exploitation and marginalization. This rural-urban migration reflects a broader pattern observed in many postcolonial societies where environmental degradation fuels socio-economic instability.

Wanja returns to Ilmorog from Nairobi, where she had sought better opportunities but faces disillusionment and hardship, including economic struggles and failed relationships. Returning to a village ravaged by environmental degradation and economic decline, she establishes a brothel to survive. This act highlights how the intersection of environmental and social injustices drives people to desperate measures. Munira captures the moral decline in Ilmorog, stating that "[t]he land that had nurtured their souls was now a desolate field, mirroring their moral decay" (p. 320).

The novel exposes how colonial exploitation's legacies continue to shape the lives of individuals and communities, revealing the deep psychological and material scars left by historical oppression. Through the characters' experiences, Ngũgĩ illustrates how colonial traumas persist through generations, manifesting in social disillusionment and economic hardship. The degradation of Ilmorog's ecosystem is a powerful symbol of the broader effects of neocolonial exploitation. The destruction of land and traditional lifestyles exemplify how capitalist projects harm both the environment and the locals.

Resilience and Indigenous Ecological Knowledge

Traditional Practices

In *Petals of Blood*, Ngũgĩ explores resilience and indigenous ecological knowledge as crucial elements for the Ilmorog community. These themes are interwoven with the villagers' struggle against the forces of neo-colonialism, showcasing how traditional practices and ecological wisdom sustain them. Ngũgĩ illustrates how Ilmorog villagers rely on indigenous ecological knowledge to survive, which serves as a form of resistance against capitalist exploitation. These practices sustain the community physically and foster social solidarity and resilience in the face of socio-environmental degradation.

One of the central traditional practices is communal farming, where the villagers work the land together to ensure communal benefit from the harvest. This collective system embodies a sustainable land-use approach. However, the villagers' connection to land is threatened when outsiders introduce commercial agriculture and privatized land ownership. This shift is signified by the arrival of multinational corporations and local elites like Kimeria, who profit from exploitative land deals. The disruption of communal farming serves as a metaphor for the disintegration of the community itself: "Our land is now fenced by the rich from the city. Where once we planted our food, now they build factories and offices" (p. 301).

The villagers also attend rainmaking ceremonies during a severe drought, showcasing their reliance on indigenous spiritual and ecological knowledge. Nyakinyua, a village elder, leads these rituals to invoke the gods to bring rain and restore the fertility of the land: "Nyakinyua took the gourd of beer and poured some onto the earth... invoking the ancestors to bring back the rain, to save their children from the drought" (p. 113). This practice highlights their deep spiritual connection to nature, contrasting with the modern capitalist mindset, which prioritizes economic growth over ecological balance. The rainmaking ceremony exemplifies how the villagers try to maintain their cultural identity and connection to land in the face of environmental degradation.

Indigenous medicinal knowledge is also crucial in sustaining the community's health. Village elders use traditional herbal remedies to treat illnesses, reflecting their understanding of the local environment and its natural resources.

This traditional healing contrasts with the alienating effects of Western medicine and commercial healthcare. When Wanja falls ill, Nyakinyua applies traditional remedies, using herbs gathered from the local environment: "Nyakinyua brought out herbs she had boiled earlier and made Wanja drink... she knew what roots healed which ailments" (p. 273). This demonstrates the villagers' reliance on indigenous knowledge for survival and well-being, even as the modern world seeks to marginalize such practices.

Lastly, cooperative labor or *Harambee* is integral to the community's resilience. In Ilmorog, villagers come together to work on each other's farms, build homes, and help during times of crisis, fostering social cohesion and mutual aid. As one villager reflects, "[i]t was only when we worked together that we survived the harshest times" (p. 159). This tradition contrasts sharply with the capitalist system that values individualism and competition. The destruction of communal spaces and shifting toward privatization under capitalist pressures erode this sense of collective identity, leading to environmental and social disintegration.

Through such traditional practices Ngũgĩ underscores the importance of indigenous ecological knowledge in sustaining the environment and the social fabric of Ilmorog. These practices stand in stark opposition to the forces of neocolonialism and capitalism, which seek to exploit the land and the people for profit.

Nyakinyua's Resilience

Resilience is depicted through the villagers' ability to adapt to environmental and socio-economic changes. Nyakinyua, one of the elders in Ilmorog, plays a pivotal role in guiding her community through environmental and socio-economic challenges. Her resilience, grounded in her profound knowledge of traditional practices, allows the community to navigate challenging times while maintaining cultural identity and solidarity.

Nyakinyua guides her community through her agricultural wisdom, reflecting her understanding of sustainable farming practices passed down through generations. She frequently reminds the villagers of their traditional ways of tending to the land, emphasizing the importance of working together and nurturing the environment. For instance, when Ilmorog faces a drought, she encourages the villagers to continue tending to their farms despite the increasingly harsh conditions: "We must still till the land as our ancestors did, for the rains will come. They always come. The land has been faithful to us, and we must be faithful to it" (p. 110). Nyakinyua's advice reflects the indigenous belief that humans must remain connected to the land and its cycles, even when external forces threaten that bond.

Her spiritual leadership is also crucial during times of environmental crisis. When the drought worsens, and the villagers feel hopeless, Nyakinyua leads the rainmaking ceremony— a traditional ritual calling upon the ancestors to restore balance. She understands that this ritual is not merely symbolic but a way to unite the community in a shared act of cultural preservation. During the ceremony, she pours beer onto the earth, offering it to the ancestors: "We pour this beer to the earth, to our ancestors, to the spirits who guard this land... Bring back the rain, bring back life to Ilmorog" (p. 113). Her leadership during this ceremony manifests how indigenous resilience draws on collective memory and cultural continuity to maintain social cohesion (Kirmayer et al., 2011, p. 87).

Additionally, Nyakinyua serves as a repository of indigenous knowledge, passing down medicinal and healing practices that have sustained the community for generations. She knows which plants can heal specific ailments. Her ability to cure Wanja reflects her deep connection to the local ecosystem: “Nyakinyua boiled roots she had gathered from the forest, herbs whose names had long been forgotten by most of the villagers. She made Wanja drink the bitter brew, chanting softly as she did so” (273). Nyakinyua’s role as a healer demonstrates how her knowledge of the environment and its resources is vital to the community’s resilience.

Lastly, Nyakinyua embodies resilience through her ability to adapt to socio-economic changes without losing her sense of self and culture. When multinational corporations threaten to displace the villagers from their land, she does not give in to despair. Instead, she encourages the community to unite and resist these external forces. Her wisdom offers guidance that strengthens the villagers’ resolve to fight for their land and way of life: “They think they can come here and buy us off, but this land is ours. We have fought for it before, and we will fight for it again” (p. 301), she declares, inspiring the villagers to defend their heritage.

Petals of Blood uses the themes of resilience and indigenous ecological knowledge to critique the exploitation of African resources and celebrate its people’s enduring spirit and wisdom. Ngũgĩ underscores the importance of these traditional practices in fostering a sustainable and cohesive community.

Ecological Resistance and Collective Action

Petals of Blood illustrates how the exploitation of natural resources by colonial and postcolonial powers has led to severe environmental degradation, which, in turn, exacerbates social inequalities. The landscape of Ilmorog, marked by drying rivers, depleted soils, and transformed lands, mirrors the socio-political turmoil experienced by its inhabitants. Eldred Jones notes that Ngũgĩ “uses the landscape of Ilmorog to reflect the disintegration brought about by the colonial encounter and subsequent neo-colonial developments” (p. 54). This section will examine two main aspects of ecological resistance in the novel: the role of collective action in confronting exploitation and the critique of Western notions of progress.

Collective Action as a Response to Exploitation

The novel presents the villagers’ journey to Nairobi as a complex symbol, representing both ecological resistance and the pursuit of better economic opportunities. The devastation of Ilmorog’s environment—marked by severe drought, the encroachment of capitalist forces, and the gradual depletion of the land—forces the villagers to confront the more profound implications of environmental exploitation. Once the community is sustained through traditional farming practices, the land is increasingly unable to support life, leading the villagers to take collective action. Jones observes that “the landscape of Ilmorog becomes a metaphor for the disintegration and decay brought about by colonial and neo-colonial policies” (p. 54). This environmental decay compels the villagers to seek intervention from the government, as they can no longer rely solely on their resources and resilience.

The villagers' journey to Nairobi, where they petition the government for assistance, is initially driven by their recognition that the ecological destruction they face is not merely natural but politically and economically motivated. The deterioration of their land is intertwined with the systemic exploitation of Kenya's natural resources by neocolonial powers and local elites. Karega highlights this fact when articulating the villagers' plight: "We can no longer survive on promises and words. We need action, we need roads, schools, water" (p. 221). Karega emphasizes the practical needs of the villagers, underscoring that their journey is as much about securing economic resources and infrastructure as it is about reclaiming their rights to a sustainable environment.

The dual purpose of the journey is further reflected in the symbolism of Nairobi itself—a city representing both the seat of political power and the promise of economic opportunity. While the villagers embark on this journey to resist the forces that have destroyed their environment, they are also motivated by the hope of gaining government support to improve their economic situation. The drought in Ilmorog ravages the land and renders traditional farming practices obsolete, leaving the villagers economically vulnerable. Therefore, their migration to the city embodies a quest for survival in the face of ecological and economic disenfranchisement.

Ngũgĩ's narrative thus intertwines environmental degradation with the socio-economic realities of postcolonial Kenya, presenting the villagers' journey as an effort to address both. By travelling to Nairobi, the villagers symbolically and practically challenge the forces of neocolonialism that have devastated their land and marginalized their communities. Additionally, they seek tangible economic improvements, reflecting the broader need for systemic change. This journey, therefore, is not solely about reclaiming the land but also about securing the means for economic survival and justice in a world shaped by global capitalism.

Ngũgĩ uses the villagers' journey to critique the neocolonial exploitation of Kenya's people and land. The highway project, ostensibly a sign of progress, ultimately serves the interests of multinational corporations and local elites, leaving the villagers further impoverished and displaced. The novel highlights how environmental degradation and economic marginalization are inextricably linked with the villagers' fight for survival, encompassing both ecological resistance and the pursuit of economic opportunities. Their journey to Nairobi is thus a powerful symbol of this dual struggle, capturing the complexities of life in postcolonial Kenya, where the well-being of the land and the people are deeply intertwined.

Critique of Western Notions of Progress

The novel's critique of Western notions of progress is evident in the stark contrast between indigenous and modern agricultural practices. The modern techniques, often characterized by monoculture, heavy reliance on chemical fertilizers, and mechanization, lead to soil depletion, reduced biodiversity, and long-term environmental degradation. In contrast, indigenous practices emphasize crop diversity, natural pest control, and soil fertility through organic means.

Ngũgĩ critiques the imposition of modern agricultural techniques by illustrating their harmful impact on the environment and the social fabric of communities. These modern techniques are shown to deplete the soil, reduce biodiversity, and disrupt the natural balance that sustained the people of Ilmorog for generations. As Karega reflects, "[t]he land is no longer ours... We work for those who control it, those who destroyed it" (p. 260). This acknowledgement of land degradation highlights how capitalist and neocolonial forces exploit and exhaust both the environment and the people.

The villagers' return to traditional farming methods is not depicted as mere nostalgia but as an act of resistance and survival. Ngũgĩ presents these indigenous practices—emphasizing crop diversity, natural pest control, and organic soil fertility—as ecologically sustainable and socially cohesive. Characters like Nyakinyua embody this resistance by maintaining traditional methods rooted in the community's cultural identity. Her traditional beer brewing from locally sourced grains and herbs reflects this connection to the land. When Ilmorog faces ecological and economic decline, Nyakinyua advises that they "go back to the old ways" (p. 275), signaling a rejection of external, profit-driven models of development in favor of practices that ensure the long-term health of both the people and the land.

This return to traditional farming also symbolises a reclaiming of agency. It is an act that reinforces communal bonds and the villagers' autonomy over their resources. Western models of agriculture, driven by profit motives, seek to alienate people from their land, turning it into a commodity rather than a shared cultural and ecological heritage. In contrast, traditional farming methods embody a holistic view of the land as interconnected with the social and spiritual lives of the people. The emphasis on reciprocity and stewardship reinforces this perspective—values that are largely absent in the modern, capitalist approaches depicted in the novel.

Ngũgĩ's narrative ultimately critiques Western agricultural practices by juxtaposing them with indigenous methods that are ecologically sound and aligned with the villagers' cultural traditions. He calls for a development model that integrates respect for local knowledge, environmental stewardship, and social cohesion, showing that progress does not always lie in adopting foreign models but in recovering and revalorizing indigenous wisdom. The villagers' return to traditional methods is thus framed as a forward-thinking strategy that offers both environmental sustainability and a deeper, more meaningful connection to their heritage.

Hope for Regeneration and Healing

In *Petals of Blood,* Ngũgĩ offers hope amidst the destruction caused by colonialism and neocolonialism, focusing on the possibility of regeneration and healing. Despite the environmental devastation and social fragmentation, nature is portrayed as a resilient force capable of renewal. This resilience is symbolized by the return of rain to the drought-stricken village of Ilmorog, signaling ecological recovery and the potential for social and cultural revival. Ngũgĩ uses this natural rebirth to metaphorize the characters' capacity to rebuild their lives and community. For example, the villagers' return to traditional farming methods represents a conscious reclaiming of their connection to the land, reinforcing their cultural identity and autonomy. This return to sustainable practices, rooted in indigenous knowledge, becomes a form of resistance against the capitalist exploitation that has eroded the environment and their way of life. Ngũgĩ's portrayal of the land's capacity for regeneration suggests that communities can envision a future where they survive and thrive through collective action, cultural resilience, and sustainable practices. This imagery of renewal offers hope that by reconnecting with their heritage and environment, the people of Ilmorog can overcome the trauma of exploitation and forge a path toward healing and justice for future generations.

In *Petals of Blood*, regeneration is symbolically represented through the land and the people of Ilmorog. The transformation of the arid, desolate landscape into a thriving community underlines the potential for renewal. Simon Gikandi states that "Ngũgĩ uses the motif of the land to symbolize both the suffering and the potential for renewal in postcolonial Kenya" (p. 70). The land's degradation under neocolonialism mirrors the broader systemic exploitation perpetuated by capitalist forces. Ngũgĩ critiques how these forces, which succeeded colonial rule, continue to erode indigenous traditions and displace rural communities. This builds on the themes in earlier works such as *A Grain of Wheat*, where land dispossession under colonial rule symbolises the economic and psychological trauma of the Kenyan people. Ngũgĩ, however, extends this critique to include the detrimental effects of neocolonial governance, which allows commercial interests to exploit the land further, leaving communities like Ilmorog in economic despair. The Trans-Africa Highway serves as a potent symbol of this dynamic, presented not as a conduit for development but as a channel for systemic exploitation, turning the once-thriving land into an arid wasteland.

Regeneration in the novel is a passive return to a pre-colonial state and an active process of resistance and reclamation. Characters like Nyakinyua embody this resistance through their connection to the land and efforts to revive sustainable agricultural practices. Unlike earlier works such as *The River Between*, which depicted indigenous traditions as a fragile counterbalance to colonial modernity, *Petals of Blood* positions these traditions as powerful resilience tools. Nyakinyua's agricultural revival illustrates the enduring value of indigenous knowledge, suggesting that cultural roots and communal effort can combat systemic exploitation and environmental decline.

The focus on collective action marks a significant thematic shift in Ngũgĩ's work. While earlier novels often highlighted the struggles of individual characters grappling with colonial trauma, *Petals of Blood* emphasizes the role of community in achieving regeneration. The people of Ilmorog, initially fragmented and despondent, come together in solidarity, such as their pilgrimage to Nairobi. This collective action symbolises their resistance to oppression and their shared commitment to rebuilding their community. As Gikandi points out, transforming Ilmorog's barren land into a thriving agricultural space represents both the potential for renewal and the necessity of collective resilience in the face of adversity. He states that "Ngũgĩ uses the motif of the land to symbolize both the suffering and the potential for renewal in postcolonial Kenya (p. 70)

The journey to Nairobi undertaken by Munira, Karega, Wanja, and Abdulla signifies a collective quest for regeneration. Their pilgrimage is not merely a physical journey but a symbolic act of seeking justice and restoration. This quest underscores the community's resilience and determination to reclaim their agency. As they confront systemic corruption and exploitation in Nairobi, they lay the groundwork for a future where Ilmorog can heal and flourish.

Ngũgĩ emphasizes the role of collective action in the healing process. The unity and solidarity among Ilmorog's villagers are crucial for their survival and regeneration. As G. D. Killam has observed, "the collective efforts of the community in *Petals of Blood* illustrate the power of unity in overcoming adversity and paving the way for a brighter future" (p. 119). This theme is poignantly illustrated during the drought when the community comes together to dig a well, symbolizing their joint effort to secure a better future.

Karega's character epitomizes the spirit of collective action. His involvement in labor movements and advocacy for workers' rights highlight the importance of solidarity in addressing systemic injustices. Karega's vision for a just society is intrinsically linked to the notion of healing, as he believes that proper regeneration can only occur through dismantling oppressive structures and establishing equitable systems.

Wanja's Journey of Healing

Among the central characters, Wanja stands out as a figure whose personal journey encapsulates the broader themes of regeneration and healing. Her evolution from victimization to empowerment is a microcosm of the community's potential for renewal. Wanja's early life is marked by hardship and exploitation, reflecting the broader struggles of Kenyan society under colonial and neocolonial rule. Born into a humble family, Wanja's aspirations for a better life are thwarted by systemic inequalities. Her initial escape to the city, driven by the allure of modernity and economic opportunity, results in further victimization. She is exploited in various jobs, eventually becoming a barmaid, a role that exposes her to the darker sides of urban life.

Carol Sicherman has noted that Wanja's experiences in the city symbolize the betrayal of Kenyan youth by the promises of modernization and economic progress (p. 99). This period in Wanja's life is characterized by a loss of dignity and autonomy, underscoring the unjust and dehumanizing effects of neocolonial economic structures and practices. Her return to Ilmorog, marked by disillusionment, signals the beginning of her journey toward healing and empowerment.

Wanja's return to Ilmorog is a pivotal moment in her journey. Disillusioned by her urban experiences, she seeks solace and purpose in her ancestral village. This return signifies rejecting the exploitative urban environment and a desire to reconnect with her roots. As Gikandi has observed, "Wanja's return to Ilmorog represents a symbolic rebirth, a search for identity and belonging in the face of alienation" (p. 73).

Once back in Ilmorog, Wanja begins to rebuild her life by engaging in entrepreneurial activities. Her decision to open a local bar and later a shop demonstrates her resilience and resourcefulness. Although initially controversial, this entrepreneurial spirit eventually becomes a means for Wanja to reclaim her autonomy and contribute to the community's economy. Her actions highlight the potential for economic independence to foster personal and communal regeneration.

Her emergence also marks Wanja's journey as a leader within the community. Although met with resistance, her entrepreneurial ventures establish her as a significant figure in Ilmorog's socio-economic landscape. Wanja's business acumen and determination to succeed reflect her inner strength and capacity for leadership. Her role evolves from marginalized to central in Ilmorog's regeneration as she gains respect within the community.

Her leadership is further exemplified during the community's pilgrimage to Nairobi. Wanja's decision to join the delegation underscores her commitment to collective action and justice. This journey symbolises a collective quest for recognition and redress, with Wanja playing a crucial role in advocating for the community's rights. As Killam has observed, "Wanja's involvement in the collective struggle highlights the importance of female leadership in the fight for social justice and regeneration" (p. 117).

Notwithstanding her progress, Wanja's journey is fraught with challenges. Her business success attracts both admiration and resentment, reflecting the complexities of social dynamics in postcolonial Kenya. Wanja's relationships with the male protagonists—Munira, Karega, and Abdulla—further illustrate the gendered challenges she faces. Her interactions with these characters reveal the persistent patriarchal attitudes that undermine women's efforts to assert their independence.

Wanja's resilience in these challenges is a testament to her strength and determination. Her ability to navigate these obstacles and continue her pursuit of economic and social empowerment underscores the theme of healing through resilience. Ali Mazrui has noted that "Wanja's character exemplifies the struggle of African women to overcome patriarchal constraints and contribute to societal regeneration" (p. 65). Ultimately, Wanja's journey is not merely about her healing but also about her contribution to the regeneration of her village—Ilmorog. Her entrepreneurial success creates economic opportunities for others in the village, fostering a sense of hope and renewal. Wanja's resilience and leadership inspire the community to envision a future where they can thrive despite the challenges posed by neocolonial exploitation.

Wanja's story culminates in a vision of a regenerated Ilmorog, where traditional values and modern aspirations coexist harmoniously. Her journey symbolises the possibility of healing by recreating identity, empowerment, and collective action. Wanja's transformation catalyzes broader societal renewal as the community rallies around her.

Wanja's journey is a profound narrative of healing and empowerment. Her evolution from a victim of exploitation to a leader and entrepreneur embodies the themes of resilience and regeneration central to Ngũgĩ's critique of postcolonial Kenya. Through Wanja's character, the novel highlights the potential for personal and communal development and renewal in the face of systemic oppression. Her story serves as a testament to the enduring spirit of Kenyan women and their vital role in the regeneration of their communities.

Conclusion

In conclusion, *Petals of Blood* provides an in-depth examination of how the enduring impacts of colonialism perpetuate suffering across generations and contribute to environmental injustice. The transformation of Ilmorog from a vibrant community to a site of ecological devastation illustrates the far-reaching effects of neocolonial exploitation, which exemplifies the indirect control and exploitation of a nation or region by a more powerful foreign entity. Environmental decline is intricately tied to the economic and social shifts brought about by global capitalism, highlighting how the depletion of natural resources exacerbates existing inequalities and perpetuates cycles of hardship and poverty. The trauma experienced by characters like Abdulla and Wanja reflects the broader psychological impact of colonial and post-colonial disillusionment, demonstrating how the historical patterns of violence are inherited through generations, shaping individual identity, agency, and economic stability. Nonetheless, their perseverance in the face of adversity attests to the resilience of the human spirit, serving as an inspiration for readers to believe in potential change.

Alongside condemning the betrayal of independence, Ngũgĩ's narrative highlights how the promises of freedom and progress often result in the ongoing marginalization and disenfranchisement of the very communities that sought liberation. The exploitation of Ilmorog's resources, coupled with the suffering endured by its residents, illustrates a larger framework where environmental and social injustices are intertwined. Recognizing this interconnectedness can illuminate the array of challenges that require attention. The novel powerfully depicts resistance, showcasing the community's struggle for justice and survival, which reflects the ability of collective action to confront and resolve these deep-rooted problems, invoking a sense of empowerment and the potential for change in its readers.

Petals of Blood ultimately offers a critical analysis of the shortcomings of post-colonial power, emphasizing the pressing need for a holistic approach to justice that recognizes and confronts past traumas and environmental deterioration. This call for a comprehensive approach to justice can engage readers and inspire them to commit to the cause. Ngũgĩ's work prompts readers to contemplate the intricate interaction between past historical influences and present-day difficulties, promoting a revolutionary perspective to restore the land and its inhabitants. An essential contribution to social and environmental justice debates, the novel emphasizes the need to address and correct the long-lasting effects of colonial exploitation.

Disclosure statement: The authors report that there are no competing interests to declare.

References

Berkes, F, and M. K. Berkes (2007). Ecological complexity, resilience, and social institutions: Cases from the developing world. *Complexity*, *13*(3), 6-10.

Breidlid, A. (2002). *Resistance and consciousness in Kenya and South Africa: Subalternity and representation in the novels of Ngũgĩ wa Thiong'o and Alex La Guma*. Peter Lang.

Fanon, F. (2004). *The wretched of the earth*. Grove Press.

Gikandi, S. (1989). Ngũgĩ wa Thiong'o and the writing of Kenyan history. *Research in African Literatures*, *20*(3), 65–82. http://www.jstor.org/stable/3819170

Gikandi, S. (2001). *Ngũgĩ wa Thiong'o*. Cambridge University Press.

Hoppers, C. A. Odora. (2002). *Indigenous knowledge and the integration of knowledge systems: Towards a philosophy of articulation*. New Africa Books.

Jones, E. (1980). The African novel in the contemporary world. *African Literature Today*, *10*, 49–63.

Killam, G. D. (1980). *An introduction to the writings of Ngũgĩ*. Heinemann Educational.

Kirmayer, L. J., Dandeneau, S., Marshall, E., Phillips, M. K., & Williamson, K. J. (2011). Rethinking resilience from indigenous perspectives. *Canadian Journal of Psychiatry*, *56*(2), 84–91. https://doi.org/10.1177/070674371105600203

Mugo, M. G. (1978). *Visions of Africa: The fiction of Ngũgĩ wa Thiong'o*. Three Continents Press.

Ngũgĩ w. T. (2005). *Petals of blood*. Penguin Books.

Shiva, V. (2010). *Staying alive: Women, ecology, and development*. Zed Books.

Authors' Biography

Hossein Nazari holds a PhD in English from the University of Canterbury. He is currently an Assistant Professor of English Literature at the University of Tehran, where he teaches Literature and translation courses. Some of his areas of research are Postcolonialism, Post-9/11 Literature, (neo/auto) Orientalism, Diaspora Studies, Women's/Gender Studies, and Travel Studies. Nazari is the winner of the Young Assistant Professors Award (Iran's National Elites Foundation) and the First Research Prize, in the Fourteenth National Research and Technology Festival of the National Library of Iran (2024) for his monograph, *Representing Post-Revolutionary Iran: Captivity, Neo-Orientalism, and Resistance in Iranian–American Life Writing* (Bloomsbury, 2022).

Reyhaneh Goudarzi holds a Master of Arts degree in English Literature from the University of Tehran. Her scholarly interests are wide and varied and include Postcolonialism, Environmental Humanities, Literary Criticism, Ecocriticism, and Subaltern Studies.

Volume 14, Issue 1
Winter and Spring, 2026
pp. 265-282

Journal of Applied Linguistics and Applied Literature: Dynamics and Advances

The Dialectics of Control: A New Historicist Excavation of Anthony Burgess's *A Clockwork Orange*

Pouyan Rezapour*

**Corresponding Author: Assistant Professor of English Language and Literature, Department of English, Faculty of Literature and Humanities, Lorestan University, Khorramabad, Iran*
ORCID: 0000-0003-0133-6932
Email: Rezapour.p@lu.ac.ir

Abstract

Emerging to challenge formalist and traditional historical criticism, New Historicism demands that literary works be analyzed within their complex socio-historical contexts, rejecting the notion of the autonomous aesthetic text. This article employs this interdisciplinary lens to excavate Anthony Burgess's controversial 1962 novel, A Clockwork Orange. It argues that the novel's dystopian vision of a youth culture addicted to ultraviolence and a state deploying oppressive behavioral conditioning functions not as mere speculative fantasy but as a critical interrogation of post-war Britain's deepest anxieties. By situating Burgess's work within the precise historical crucible of the Cold War, burgeoning consumer capitalism, a crisis in political authority, and the rising influence of Skinnerian behaviorism, this analysis demonstrates how the novel refracts the period's pervasive cultural tensions. Ultimately, this study contends that A Clockwork Orange is a seminal document of its time, a symbolic act that absorbs and artistically transforms the very real discourses of social control, moral panic, and ideological conflict that defined the early 1960s.

Keywords: A Clockwork Orange, Anthony Burgess, Behaviorism, New Historicism, Post-war Britain

ARTICLE INFO
Research Article
Received: Thursday, September 11, 2025
Accepted: Sunday, February 1, 2026
Published: Wedensday, April 1, 2026
Available Online: Sunday, February 1, 2026
DOI: https://doi.org/ 10.22049/jalda.2026.30990.1878

Online ISSN: 2821-0204; Print ISSN: 28208986

Introduction: Contextualizing the Ultraviolence

The publication of Anthony Burgess's *A Clockwork Orange* in 1962 sent seismic shockwaves through the literary establishment and the public consciousness. Its unflinching graphic depictions of "ultraviolence," sexual assault, and a profoundly nihilistic youth culture, all narrated in a bizarre Anglo-Russian slang lexicon dubbed "Nadsat," ensured its immediate and enduring notoriety. The novel's confrontational style provoked intense debate, polarizing early critical reception. Much of the initial scholarship focused on its sensationalistic surface, engaging in contentious debates over its morality, its potential to corrupt, and its very artistic merit. As the initial shock subsided, later interpretations leaned towards more philosophical and theological explorations, grappling with its profound questions concerning free will, ethical choice, and theodicy, which is the question of why evil exists in a world governed by a benevolent God. However, while these readings are valuable, to categorize the novel solely as a transgressive shocker or an abstract moral-philosophical fable is to critically overlook its profound and intricate embeddedness within the specific historical moment of its creation. The emergence of New Historicism as a dominant critical paradigm in the late twentieth century provides the necessary theoretical tools to unearth this essential context, arguing persuasively that a literary text is never an autonomous aesthetic object but an active participant in the complex circulation of what Stephen Greenblatt termed the "social energy" of its era.

The historical context of a literary work is the very bedrock of the New Historicist project. This approach fundamentally redefined the relationship between text and context, distinguishing itself from traditional historical scholarship. As Cox and Reynolds adeptly define it, "new historicism can be distinguished from 'old' historicism by its lack of faith in 'objectivity' and 'permanence' and its stress not upon the direct recreation of the past, but rather the processes by which the past is constructed or invented" (1993, p. 4). This epistemological shift is profound. New Historicists, influenced by poststructuralist thought, are ardent believers in the textuality of history. They operate on the premise that the studying of history is unable to bring an objective, full truth to light precisely because the historical record is itself a narrative, a constructed text. As M. Keith Booker asserts in *A Practical Introduction to Literary Theory and Criticism*, "all historical accounts are partial, and all are conditioned by the terms of inquiry" (1996, p. 135). Therefore, the New Historicist study of literary works is pivoted not on a naive reconstruction of the past, but on the scrupulous study of the discourses and power structures that constituted the dominant culture of the period of the text's production.

This methodology inherently acknowledges significant challenges. The scholar's task is fraught with complexities; the study of a present culture is immense, and the study of past cultures is even more so, mediated by documents that are often incomplete, unreliable, or filtered through the biases of those who preserved them. Furthermore, every culture is a vast and contradictory network of practices and beliefs, meaning any analysis will inevitably result in the biased inclusion and omission of certain information. The influence of the materials scholars choose to include “have a powerful impact on the historical visions they produce” (Booker; 1996, p. 136).

Crucially, scholars must be reflexively aware that their own social standing and ideological position in the present culture “exerts a powerful influence on the ways they view cultures from the past” (Booker, 1996, p. 136). It can be concluded that a complete and whole picture of a past culture is an impossibility; the New Historicist can only arrive at a fragmentary vision, a partial construction realized through the lens of their own contemporary concerns.

Anthony Burgess’s novel is a prime candidate for such an analysis precisely because of the vehement cultural reaction it ensued. The novel provoked havoc and profound discontent from critics and readers alike, a controversy that stemmed directly from its explicit depiction of gruesome, amoral violence. The figure of Alex and his teenage droogs, committing ghastly crimes without any pangs of conscience, represented a terrifying social threat. Contemporary society, in the early 1960s, was largely unwilling to accept such vivid scenes of violent criminality being propagated by a literary novel. The problem became further complicated and the cultural discourse amplified when, nine years later, the famed director Stanley Kubrick directed and produced a seminal film adaptation. The movie’s success was arguably even greater than the novel's, reaching a mass global audience. This fame meant the story’s provocative ideas and images were disseminated more widely, which in turn generated a larger and more vocal number of opponents to its broadcasting, leading to its controversial withdrawal in the UK by Kubrick himself.

Burgess’s narrative depicts a society brimming with violence and crime emanating from its youth, matched only by an unreasonable and equally brutal response from the authorities and the structures of power, including the ineffectual parents, the cynical officers of the law, the expedient court, the compromised Church, and the totalitarian government. This dual critique suggests a deep engagement with the failures of both social anarchy and state control. Henceforth, the central intention of the present study is to analyze the novel from a rigorous New Historicist perspective. This entails examining the prevalent culture and society of post-war Britain, its evolving norms and taboos, and the officials and executives in power across the spheres of politics, economics, and religion. Specifically, the youth culture arising in the aftermath of the World Wars and during the peak of the Cold War will be analyzed as a key historical formation. These elements, it is argued, were all employed by Burgess, both consciously and unconsciously, and each comes to have a significant degree of explanatory power regarding the social, historical, and political standards and prohibitions that the novel simultaneously reflects and critiques. Through this lens, *A Clockwork Orange* is revealed not as a prophecy of a distant future, but as a brilliant and dark refraction of the very real cultural tensions of its own time.

Methodology

This study employs a qualitative, hermeneutic methodology grounded in the principles of New Historicist criticism. Its primary intervention is the systematic application of this framework to Anthony Burgess's *A Clockwork Orange* (1962). The core objective is to demonstrate that the novel functions not as a speculative prophecy but as a seminal document of its time, a symbolic act that absorbs and artistically transforms the precise discourses of social control, moral panic, and ideological conflict that defined post-war Britain. The methodology is fundamentally text-centered, treating the novel's specific scenes, linguistic innovations, and character dynamics as primary evidence from which to excavate historical context.

The analysis is constructed through a sustained engagement with primary sources, Burgess's novel and his ancillary writings, and a wide range of secondary sources drawn from peer-reviewed academic journals and scholarly monographs on New Historicist theory, dystopian literature, and post-war British culture. Operationally, this involves a "thick description" of the literary text, reading it alongside non-literary "texts" and discourses from its historical moment e.g., political speeches, psychological studies, media reports. This practice is guided by Louis Montrose's foundational dialectic of "the historicity of texts and the textuality of history," which posits that the literary work is both shaped by and actively shapes the network of social energy in which it circulates. By situating the novel within the precise historical crucible of the early 1960s, this methodology seeks to illuminate how *A Clockwork Orange* critically interrogates the period's deepest anxieties concerning youth, state power, behavioral science, and consumer capitalism.

Operationalizing the Theory: A Schematic Approach

To translate the principles of New Historicism into a concrete analytical practice, this study employs a specific methodological schema. Each core tenet of the theory directly informs a corresponding mode of engagement with *A Clockwork Orange*, ensuring the analysis remains a disciplined excavation of historical discourse rather than a general contextualization. The following schematic outlines this operational logic:

The foundational premise of the historicity of texts guides the investigation's primary orientation. This principle demands that the novel be read not as a timeless philosophical allegory but as an artifact saturated by the specific anxieties of its post-war British moment. In practical terms, this means interpreting the pervasive "ultraviolence" of Alex and his droogs as a direct literary refraction of the era's very real moral panic over juvenile delinquency, embodied by the Teddy Boys and Mods and Rockers, and the pervasive media discourse around a destabilizing "youth question."

Conversely, the principle of the textuality of history dictates the analytical method. It acknowledges that our access to the past is always mediated by surviving narratives and documents. The methodology, therefore, involves reading Burgess's literary text alongside key non-literary "texts" from the early 1960s. For example, the novel's Ludovico Technique is analyzed not in a vacuum but is placed in dialogue with contemporaneous discourses: B.F. Skinner's polemics on behavioral technology, parliamentary debates on crime and punishment, and journalistic reports on aversion therapy. This intertextual practice reveals the novel's satirical engagement with a precise scientific and political conversation.

The New Historicist focus on the circulation of social energy shifts the analysis from static reflection to dynamic cultural participation. It prompts an examination of how the novel itself entered and altered the discursive field. This involves tracing the reception of the work, particularly after Stanley Kubrick's 1971 film adaptation, when real acts of violence were sensationalized in the press as "Clockwork Crimes." This phenomenon demonstrates how the artwork became an active agent, a new lens through which societal fears were both expressed and amplified.

The analysis of power and social control, deeply informed by Foucault, provides the central thematic lens. It directs the critique toward the mechanisms of discipline employed by various institutions within and against the novel's world. The investigation interrogates how both Alex's gang, through brute force, and the state, through the Ludovico Technique, seek to impose order, creating a critical dialectic. The state's project of rendering Alex a predictable "clockwork orange" is thus revealed not as a solution to social chaos, but as its tyrannical mirror image, two sides of the same coin of dehumanizing control.

Finally, a nuanced consideration of the author as a nexus of historical forces enriches the reading without resorting to biographical reductionism. It contextualizes Burgess's creative choices within the confluence of his personal experiences, such as the traumatic assault on his first wife, his Catholic education, and his colonial service in Malaya, and the broader currents of post-war disenchantment. These experiences form the personal substrate through which the era's larger anxieties about violence, free will, and cultural hybridity were filtered and artistically transformed in the invention of Nadsat and the novel's moral architecture.

Literature Review

A comprehensive review of the scholarship confirms that while *A Clockwork Orange* has been analyzed from a multitude of critical perspectives, a dedicated and sustained New Historicist reading remains a significant, though not entirely unaddressed, gap in the literature. It is crucial to situate this study within the broader critical conversation.

Early scholarship established key avenues of inquiry that remain influential. Robert O. Evans's (1971) foundational linguistic analysis of Nadsat established the novel's stylistic innovation as a critical site of meaning. Philosophical and ethical readings, like those of Bobby Newman (1991), rigorously examined its engagement with Skinnerian behaviorism, often framing the central conflict as one between free will and determinism. Subsequent studies employed adjacent theoretical lenses; for instance, Pat Gehrke's (2001) Foucauldian critique effectively analyzes the construction of criminological subjectivity, while Lauren Menking (2012) offers a comparative study of autonomy.

The evolution of scholarship in the past decade indicates a strong trend towards more culturally and historically situated readings, yet often without fully embracing the integrative apparatus of New Historicism. For example, recent work by Böll (2022) precisely historicizes the character of Alex within the post-war British media's construction of the "juvenile delinquent," tying the novel's ultraviolence directly to the era's specific moral panics. Important meta-critical work by Van Lier (2018) on the critical reception shaped by the omitted twenty-first chapter highlights how the novel's own material and publication history is a constitutive part of its cultural meaning, a concern central to New Historicist analysis.

While New Historicist readings remain scarce, significant scholarly work has positioned *A Clockwork Orange* as a critical intervention in post-war British political discourse. Scholars have convincingly analyzed the novel's engagement with the period's defining anxieties concerning state power and social control. Jackson (2018) provides a philosophical grounding for these analyses, arguing that the novel's central conflict exposes the mechanisms of what he terms the "technology of behavior," a fusion of state power and behavioral science aimed at total social control. Together, these studies establish a firm foundation for understanding the novel's political dimensions, yet they stop short of the comprehensive, discourse-driven synthesis that a full New Historicist analysis would undertake.

Theoretical Foundations: The New Historicist Project

Defining the New Historicist Enterprise

This analysis is grounded in the principles of New Historicism, a critical practice that reconceives the relationship between literary texts and history. As articulated by its key proponents, this approach provides the essential framework for reading *A Clockwork Orange* not as a timeless allegory but as a dynamic participant in the cultural negotiations of its historical moment.

The methodological cornerstone is Louis Montrose's formulation of "the historicity of texts and the textuality of history" (1989, p. 20). This dialectic insists on two inseparable premises: first, that a literary work is a culturally embedded artifact, saturated by the conditions of its production, its historicity; and second, that our access to the past is always mediated by textual traces, archives, documents, narratives, that are themselves constructs shaped by power, history's textuality. Consequently, the critic's task is to trace the circulation of social energy, a term Stephen Greenblatt uses to describe how power, anxieties, and pleasures are negotiated and transferred between a literary text and the broader network of contemporary discourses (Greenblatt, 1989).

To execute this task, the analysis employs two key instruments. From anthropologist Clifford Geertz, it adopts the technique of "thick description" (1973), the meticulous interpretation of specific social actions (or, by extension, literary scenes) to unpack the layered webs of cultural meaning within which they are situated. From Michel Foucault, it draws the analytical focus on discursive practices and diffuse power relations (1977). Foucault's work demonstrates how institutions produce knowledge to categorize, discipline, and normalize subjects, a model crucial for analyzing both the Ludovico Technique's behavioral conditioning and the state's mechanisms of control in Burgess's novel.

Therefore, this study reads *A Clockwork Orange* as one "text" among many in the early 1960s discursive field. It performs thick descriptions of the novel's key scenes, the Korova Milkbar, the Ludovico sessions, the domestic spaces, and reads them against contemporaneous non-literary discourses on youth delinquency, consumer affluence, behaviorist psychology, and Cold War politics. The goal is to illuminate how Burgess's novel absorbed, concentrated, and artistically transformed these specific historical energies.

Foundational Influences: Foucault and the Archeology of Power

The single most profound influence on New Historicist thought is the work of Michel Foucault. His seminal studies, including *Madness and Civilization* (1961), *Discipline and Punish* (1975), and *The History of Sexuality* (1976), provide the philosophical and methodological blueprint for the movement. Foucault's historical investigations analyze how societal institutions—prisons, clinics, asylums, and schools—function as discursive "apparatuses" that produce knowledge and exert power through systems of surveillance, normalization, and discipline. He argues that society is not a monolithic structure but a complex and often contradictory amalgamation of interrelating discourses rooted in power. His methodology involves tracing how these discourses constitute subjects and define deviance.

For New Historicists, Foucault's work is indispensable for several reasons. It displaces the traditional Marxist focus on economics with a more diffuse model of power that operates at the capillary level of everyday life. It demonstrates that power is not merely repressive but also productive, creating categories of knowledge and social norms. Furthermore, Foucault's rejection of grand historical narratives or "great evolutionary process[es]" (Foucault, 1977, p. 129) in favor of analyzing specific, local "discursive practices" licenses New Historicists to draw connections between a literary text and a seemingly unrelated non-literary document, such as a colonial report or a medical journal. Crucially, Foucault was reflexively aware that "his own habits of thought shape the outcome of these investigations" (Booker, 1996, p. 137), a self-awareness that New Historicists adopt in their own critical practice.

Anthropological Insights: Geertz and Thick Description

If Foucault provides the theory of power, the cultural anthropologist Clifford Geertz provides the primary methodological tool for its application. Geertz's definition of culture as "an historically transmitted pattern of meanings embodied in symbols, a system of inherited conception expressed in symbolic forms by means of which men can communicate, perpetuate, and develop their knowledge about and attitudes toward life" (1973, p. 89) shifted the anthropological goal from establishing universal laws to interpreting specific cultural systems. His celebrated technique of "thick description" involves the meticulous analysis of a culture's smallest and most local practices, such as a Balinese cockfight, to unravel the complex webs of significance that constitute social life.

Geertz's influence on New Historicism is profound. His methodology legitimized the practice of treating a literary text not as a privileged, autonomous object but as one cultural artifact among many, to be read alongside and against contemporaneous non-literary "texts"—such as legal documents, travelogues, diaries, or political pamphlets. The attention he paid to marginal or peripheral cultural practices encouraged New Historicists to look beyond the literary canon and to question the hierarchies of value embedded within traditional historiography. By focusing on symbolic actions, Geertz provided a model for how to read a novel like *A Clockwork Orange* not just for its plot, but as a symbolic act that participates in the broader cultural conversations of its era.

Key Principles and Critical Practices

The synthesis of Foucault's power analytics and Geertz's thick description results in a distinct style of criticism that redefines the relationship between literature and society. It explicitly rejects the New Critical dogma of the text as a self-contained verbal icon, insisting instead that a literary work is produced at the intersection of the author's interaction with the world and is interpreted by critics who bring their own historical situatedness to the process. H. Aram Veeser (1989) usefully distills the movement's key premises into five central claims, which are as follows. Initially Veeser points out that every expressive act is embedded in a network of material practices. Then, that every act of critique or opposition inevitably uses the tools of the system it condemns and risks complicity. In third place Veeser claims that literary and non-literary texts circulate inseparably and must be read intertextually. The fourth claim is that no discourse, whether artistic or documentary, offers access to unchanging truths or a stable human nature. And finally, that the critic's own language and method are implicated in the economy they describe.

In practice, this leads to the kind of work pioneered by Stephen Greenblatt in Renaissance Self-Fashioning (1980), which demonstrates how the identities fashioned by Renaissance authors were constrained and enabled by powerful institutions like the church, court, or colonial administration. The success of this approach has led to its widespread application across literary periods, from Walter Benn Michaels's work on American naturalism to Vincent Pecora's studies of modernism. Its very popularity, however, has sparked internal critique, as noted by Montrose (1986), who warned against the movement hardening into a new orthodoxy, and who reminds practitioners that the academy is a site of ideological contestation where scholars must be conscious of their own power and positionality (1989, p. 416). This self-reflexive, politically engaged, and interdisciplinary stance remains the enduring legacy of the New Historicist project.

Applying the Lens: From Theory to Textual Practice

The principles outlined above are not applied as a rigid template but as an interpretive lens. The practical methodology of this analysis, therefore, will be to allow the novel's specific textual details, its scenes, linguistic innovations, and character dynamics, to generate the central analytical questions. For instance, the clinical brutality of the Ludovico Technique sessions does not merely "illustrate" Foucaultian discipline; it actively stages it, allowing us to ask how Burgess's narrative form makes the abstract "technology of behavior" viscerally real. This text-centered approach ensures that our New Historicist reading remains grounded in the literary artifact itself, using its symbolic constructions as the primary evidence for its historical engagements.

A New Historicist Reading: Post-War Britain Refracted in Dystopian Form

A Methodology of Textual Anchoring

A New Historicist reading posits the literary text as a cultural artifact. To operationalize this, the following analysis will proceed not by layering historical context onto the novel but by excavating context *from* the novel's own symbolic fabric. Each section will originate in a close examination of a pivotal textual element—the dysfunctional domestic space of the F. Alexander home, the synthetic consumerism of the Korova Milkbar, the hybrid slang of Nadsat. These elements serve as our primary archives. We will subject them to a "thick description," asking what specific historical anxieties, discourses, and power relations are encoded within their narrative presentation. This method inverts the casual, illustrative reference; here, the text is the constant nucleus, and historical context forms the explanatory field that radiates from it. Through this practice, *A Clockwork Orange* is revealed not as a work *about* its era, but as a performative act *within* it.

Key Analysis: The Theory-Practice Dialectic in Action

The true strength of a New Historicist approach lies in the dynamic, dialectical relationship between its theoretical principles and its practical methodology. This is not a one-way application of theory to text, but a continuous loop where each informs and refines the other.

From Theory to Text (Practical Application): The core principle of the "textuality of history" is not merely an abstract philosophical stance. It actively guides the practical research method of placing Burgess's invented Nadsat slang, a fusion of Russian roots with English, alongside Cold War political discourse from the early 1960s. This practice reveals how the novel linguistically embodies the era's paranoia about ideological infiltration and the blurring of East-West boundaries, showing history's "texts" embedded in the novel's very language.

From Text to Theory (Refinement & Challenge): Conversely, a close reading of the novel's material history and divergent endings challenges and refines theoretical assumptions. The fact of the omitted 21st chapter in the early American edition and Kubrick's adaptation forces the critic to confront how a text's material production, marketing, and adaptation shape its public "historicity." This practical observation leads back to theory, highlighting how market forces and cultural reception can alter a work's intended intervention, a nuance central to New Historicism's focus on the circulation of cultural energy.

Beyond Reflection to Active Interrogation: A simplistic historicist practice might claim the novel "reflects" 1960s fears of youth violence. A sophisticated New Historicist practice, deeply informed by its theoretical commitment to analyzing power, demonstrates that the novel actively participates in and interrogates the cultural conversation. It does not simply mirror a moral panic; it uses satire to stage a dialectic of control, critiquing both the anarchic, consumerist violence of Alex and the oppressive, behaviorist response of the state. The text becomes a site where the era's competing discourses of individual freedom and social order clash.

The Political Context: A Realm in Decline

The novel's dystopian landscape is fundamentally shaped by a crisis of institutional legitimacy, directly refracting post-war Britain's loss of faith in authority. Burgess excavates this anxiety not through grand declarations, but through the granular failure of every entity meant to provide order and care, from the state to the family.

The most potent symbol of this failure is P.R. Deltoid, the Post-Corrective Advisor. His interactions with Alex reveal a state apparatus concerned with bureaucratic management, not rehabilitation. After Alex's arrest, Deltoid's visit is devoid of empathy, centering instead on his own professional embarrassment, "'You've made me look a right govoreet,' he said.'I was your friend, Alex, and you let me down'" (Burgess, 1962, p. 68). The term "friend" is bitterly ironic, exposing the state's paternalistic role as a hollow performance. His function is not to correct but to label and dispose, highlighting an institution that has abandoned its moral purpose.

This institutional bankruptcy is mirrored and amplified in the domestic sphere, the traditional foundation of social authority. Alex's parents are portrayed as profoundly passive and ineffectual. Upon his return from prison, his father's welcome is weak and mediated by tabloid discourse: "'We've got your room ready... We read about what they did to you... It was in all the papers'" (p. 137). Their concern is derivative and passive, and their authority is so nullified that they are quickly displaced by a lodger, Joe. This signals the home's transformation from a sanctuary into a negotiable economic asset, underscoring the collapse of paternal authority.

This comprehensive vacuum of legitimate authority, from family to state, creates the essential conditions for the novel's central dialectic. The narrative portrays a society in the throes of what might be termed a post-ideological collapse, a condition specific to post-imperial Britain in the wake of the Suez Crisis (1956). With the grand narratives of progress, order, and paternalistic authority shattered, a profound void opens at the center of civic life. Into this void steps first the anarchic, tribal authority of Alex's gang, and subsequently, the state's brutal, technocratic solution in the Ludovico Technique. The novel thus presents a stark diagnosis of a national crisis: totalitarian control and random ultraviolence are not opposites but symbiotic pathologies, both born from the same foundational collapse of trusted institutions. The narrative argues that the vacuum left by crumbling traditional authority is inevitably filled by either anarchic violence or cynical state control, offering no viable alternative.

The Economic Context: The Affluent Society and its Discontents

Burgess's genius lies in portraying Alex's ultraviolence not as a result of poverty, but as a grotesque extension of consumer affluence and the commodification of experience. In the world of the novel, violence becomes the ultimate consumer good, a high-stakes leisure activity for bored youths in a sensation-seeking economy. This critique is anchored in the iconic setting of the Korova Milkbar, a site that offers a "thick description" of this new societal logic. It is not a place for communal bonding but a temple to synthetic, individual consumption, where clients drink "moloko plus" (milk laced with drugs) served by "devotchkas" amidst "rotten" statuary furniture. This environment does not satisfy a natural need; it manufactures artificial desires for chemical stimulation and aestheticized novelty, modeling a culture where authentic feeling is systematically replaced.

Alex and his droogs are the perfect products of this affluent society. Their crimes consistently lack economic motive, transforming violence into a consumable experience pursued for its own sake. The invasion of the writer F. Alexander's home is a key example; they seek not valuables but the experiential "kick" of terror and dominance. Alex narrates the assault with chilling aesthetic detachment, "Then there was the quiet, and we were breathing like asthmatic, and then we cracked into lovely laughter" (Burgess, 1962, p. 23), framing brutality as a form of performance art. This presents violence as 'aestheticized leisure,' a logical endpoint of an affluent society that treats all experience as a consumable

The novel's damning critique, therefore, is twofold. First, it dramatizes the spiritual emptiness of the "affluent society" diagnosed by Galbraith (1958), showing how material plenty can generate a violent, nihilistic search for authentic feeling. Second, it reveals a terrifying continuity between the disease and its cure. The Ludovico Technique ironically mirrors the very consumerist logic it purports to counteract; it seeks to "fix" undesirable behavior (a broken product) through a swift, technological intervention, utterly ignoring the underlying social and spiritual malaise. In positioning both anarchic violence and authoritarian control within the same framework of empty, mechanistic transaction, Burgess argues that a society which reduces human needs to problems of technical management is already halfway toward the dystopia it fears.

The Social Context: Moral Panic and the "Youth Question"

The most immediate and visceral social context for the novel was the pervasive moral panic over youth culture. The figure of the "Teddy Boy" in the

1950s, and later the "Mods and Rockers" of the early 1960s, became potent media symbols of social decay and moral decline. Sensationalist newspaper reports of gang violence and clashes at seaside towns like Brighton, famously analyzed by Stanley Cohen (1972) in his study of moral panics, created a public perception of a nation under threat from its own amoral, uncontrollable youth. Politicians and public figures demanded swift and effective solutions to this perceived "youth problem."

The discourse around a "youth problem" was fueled by concerns regarding the waning influence of traditional leftist politics, the ascendancy of consumer culture, and the rise of new youth subcultures (e.g., Teddy Boys, jazz clubs) that were perceived as disconnected from established political structures (1959). Prominent British New Left intellectual Stuart Hall interpreted this not as apathy but as a disenchantment with mainstream politics, arguing that youth were expressing dissent through cultural styles. Consequently, he urged political figures, particularly within the Labour Party, to recognize and constructively engage with these cultural expressions.

Burgess, who worked as a teacher, had direct experience with this generation. Alex and his droogs are a direct literary extrapolation of these media-fueled fears. Their fashion, the "cravat," the "heavies" (boots), and the "old jelly mould" (codpiece), parodies and exaggerates contemporary youth styles. Their mindless, aestheticized violence reflects the media's worst nightmares. The novel engages directly with the central tenet of the contemporary debate which can be acknowledged as the question, are these youths inherently evil, or are they products of a sick society? The state's answer, the Ludovico Technique, is a terrifyingly logical extension of the calls for a quick, technological fix to a complex social issue. It represents the ultimate paternalistic overreach, the desire to eliminate the symptom, violent behavior, without addressing any of the underlying societal causes, such as the very affluence and spiritual emptiness that the state itself promotes.

The Scientific Context: Behaviorism and the "Technology of Behavior"

The Ludovico Technique stands as one of literature's most potent critiques of behaviorist psychology, dramatizing the ethical catastrophe of reducing moral choice to mechanistic conditioning. Burgess moves beyond abstract debate, forcing the reader to experience the horror of having one's humanity engineered away. The procedure's mechanism is rendered through Alex's agonizing sensory experience: "The nausea started now... I was heaving away... and then I felt like I was drowning" (Burgess, 1962, p. 115). This clinical brutality underscores the state's reductive goal, summarized by Dr. Brodsky's amoral declaration: "We are not concerned with motive... We are concerned only with cutting down crime!" (p. 115). This stance directly engages with B.F. Skinner's (1971) contemporary arguments for a "technology of behavior," personifying the era's scientific optimism that treated the human mind as a machine to be rewired.

The true philosophical horror of the technique, however, is revealed in its catastrophic side effect. By using Beethoven's Ninth Symphony as the score to violent films, the state inadvertently conditions Alex against sublime art. His anguished cry, "It's a sin! Using Ludwig van like that!" (ibid, p. 132), exposes the treatment's fundamental barbarism: it cannot discriminate between destructive violence and transcendent beauty, erasing his capacity for deep emotional response. This narrative turn validates the Prison Chaplain's humanist and theological warning, voiced earlier: "When a man cannot choose, he ceases to be a man" (p. 93). Burgess, therefore, does not merely reference the period's debate between behaviorism and free will; he dramatizes its human cost, arguing through narrative consequence that a psychology ignoring the soul annihilates the interior self.

This literary intervention constitutes a powerful critique of the period's scientific paradigm. Burgess, through the Chaplain and the narrative's sympathy for Alex's loss, aligns with critics who saw in such science a fundamental threat to existential freedom. The Chaplain's argument that goodness is meaningless without free will provides the novel's core theological rebuttal to behaviorism, asserting that genuine morality necessitates the risk of evil choice. Thus, *A Clockwork Orange* functions as a seminal critique, insisting that any technology seeking to engineer goodness by destroying choice ultimately annihilates the very humanity it claims to protect.

The International Context: Cold War Paranoia and Linguistic Politics

The novel's most ingenious formal feature, the Anglo-Russian slang Nadsat, is far more than a stylistic coincidence. It is a profound narrative embodiment of Cold War paranoia, materializing the fear of ideological infiltration and the blurring of cultural boundaries in a nuclear age.

Nadsat linguistically fuses the enemy's lexicon with the protagonist's mind. Words like *droog* (friend, from Russian *drug*), *horrorshow* (good, from *khorosho*), and *devotchka* (girl, from *devochka*) are the basic fabric of Alex's consciousness (Burgess, 1962, p. 3). This creates a profound dissonance, the ultra-English hoodlum, symbol of domestic social decay, thinks and speaks with a Russian-derived vocabulary. This embodies the period's deepest anxiety, that the ideological conflict with the Soviet East was not just external but an internal war for the soul of Western youth.

The geopolitical tension is further reflected in the state's own methods. The Ludovico Technique, with its clinical, dehumanizing efficiency, mirrors the cold, mechanistic stereotype of Soviet-style control. The novel thereby suggests a terrifying convergence, the Western state, in its panic to quell internal disorder, adopts a form of psychological coercion that mirrors the very totalitarianism it purports to fight. Burgess thus uses language and plot to argue that the Cold War's binary "us vs. them" logic was breaking down, creating monstrous hybrids like Alex, a figure of English vandalism speaking the language of Soviet ideology.

The Authorial Context: Burgess' Biographical Refractions

A New Historicist reading must consider how Burgess's biography positioned him as a unique link through which the era's major currents filtered. His life experiences are not sources for one-to-one allegory but are the specific personal fractures through which broader historical traumas were refracted into art.

Several key biographical strands are critical. First, the brutal assault and rape of his first wife, Lynne, by American G.I.s during the WWII blackout is starkly transmuted into the novel's central traumatic event: the gang's attack on F. Alexander and his wife. This personal knowledge of random, institutionalized violence profoundly shapes the novel's cynical view of both criminal and state-sanctioned brutality. Second, his Catholic upbringing directly informs the novel's moral architecture, providing the framework for the Chaplain's arguments about soul, free will, and the necessity of choosing goodness. Third, his time as an education officer in Malaya and Brunei exposed him to the mechanics of colonialism and cultural hybridity, influences seen in the constructed language of Nadsat and the novel's preoccupation with the control of subject populations.

Furthermore, Burgess identified with the "Angry Young Men," a literary movement deeply critical of the British class system and establishment hypocrisy. This alignment fuels the novel's satirical contempt for all figures of authority, from politicians and police to liberal intellectuals like F. Alexander, whose revolutionary ideals are revealed to be as self-serving and violent as the regime he opposes. Burgess, through this biographical lens, becomes less an isolated author and more a symptomatic figure, whose personal grievances and intellectual passions channeled the postwar disenchantment, theological anxiety, and anti-authoritarian spirit of his time directly into the symbolic form of his most famous novel.

The Generic Context: Dystopian Tradition and Its Evolution

A Clockwork Orange represents a pivotal evolution in the dystopian genre, shifting the primary threat from the external coercion of a totalitarian state to the internal, psychological manipulation of the individual by a therapeutic state, while introducing spiritual emptiness as a new social danger.

Classic dystopias like George Orwell's magnum opus *1984* (1949) and Huxley's *Brave New World* (1932) focus on state control through surveillance (*Big Brother*) or biological and social engineering (the World State). Burgess's innovation is the Ludovico Technique. Its horror is intimate and psychological; the state doesn't just punish Alex's body or control his thoughts, it reprograms his very nervous system, making him a "clockwork orange", organic on the outside, mechanical within. This internalizes the dystopian terror, moving it from the prison cell to the conditioned mind. Furthermore, Burgess introduces a new enemy absent in earlier dystopias, affluent nihilism. Alex's society is not one of deprived, oppressed masses, but of bored, over-stimulated consumers. The state isn't creating his evil; it is ineptly trying to manage a symptom of its own success in creating material comfort without spiritual purpose.

Synthesis: The Novel as a Historical Refraction

In conclusion, a New Historicist reading reveals that *A Clockwork Orange* is a complex symbolic act that absorbed the central anxieties of its time and refracted them through a darkly comic and philosophically rich dystopian lens. It is a novel about the crisis of authority in a post-imperial Britain, the spiritual emptiness of a consumer society, the moral panic over rebellious youth, the ethical nightmare of behaviorist psychology, and the pervasive paranoia of the Cold War. Burgess was not a prophet predicting a distant future, he was a satirist diagnosing his own present, pushing the discourses of his era to their most logical and terrifying extremes. The novel endures because the dialectic it explores, between individual autonomy and external control, is perpetual. Our contemporary debates about cancel culture, social media algorithms, neuro-marketing, and state surveillance are all variations on this same theme, ensuring that Burgess's unsettling question remains as relevant as ever: what does it mean to be human in a world increasingly eager to engineer our humanity?

Conclusion

A New Historicist excavation of Anthony Burgess's *A Clockwork Orange* reveals a novel that is profoundly and inescapably a product of its specific historical moment. Through its darkly comic and philosophically rich dystopian lens, the novel refracts the pervasive fear of the newly empowered youth, the simultaneous allure and terror of burgeoning behavioral science, the deep-seated paranoia of the Cold War, the creeping emptiness of mass consumerism, and the palpable collapse of traditional moral and institutional authority. Burgess was less a clairvoyant than a masterful satirist, conducting a grim diagnosis of his own present. The world of Alex and his droogs is the world of the Teddy Boys and the Mods and Rockers, the chilling theories of B.F. Skinner and the brutal pragmatism of Cold War politicians, all pushed to their most logical and terrifying extreme.

The enduring power and unsettling genius of the novel lie precisely in its masterful weaving of these disparate historical threads into a coherent and devastating critique. It dismantles any system, whether it is the overt, state-sanctioned oppression of the Ludovico Technique or the violent, individualistic anarchy of Alex's gang, that seeks to deny the messy, dangerous, and fundamentally essential human capacity for moral choice. The novel posits that true humanity resides not in enforced goodness, but in the God-given, and often terrifying, freedom to choose between good and evil. This is the core of Burgess's Christian existentialist argument, framed within a narrative of ultraviolence and state control.

Crucially, *A Clockwork Orange* endures because the central dialectic it explores which is the struggle between individual autonomy and external control, is a perpetual feature of the human condition. Our contemporary debates are merely new iterations of this ancient conflict. The novel's themes find startling resonance in modern phenomena for instance, the mechanics of "cancel culture" echo the Staja's attempts to reform and ostracize, social media echo chambers and algorithmic curation of reality mirror the government's manipulation of Alex's conditioned responses, the fields of neuro-marketing and data surveillance represent a sophisticated, commercialized extension of the behaviorist principles that underpin the Ludovico Technique, and finally the ongoing tensions between state security and personal liberty continue to reflect the Cold War-era paranoia that saturates the novel's atmosphere.

Therefore, *A Clockwork Orange* remains a vital and indispensable work not merely for the philosophical question it poses, but for the formal and textual way it embodies that question's historical urgency. The novel's enduring challenge is crystallized in its own material history, the omission of the final chapter in its initial American edition. This editorial choice, privileging a cycle of unrepentant violence over Burgess's original vision of weary, potential redemption, itself becomes a New Historicist datum. It shows the mid-century market's preference for a darker, more deterministic parable over a cautiously hopeful one, reflecting a cultural moment deeply skeptical of organic change. The novel thus forces each generation to confront its central dilemma not as an abstraction, but through the very experience of reading, through the jarring music of Nadsat, the visceral horror of the conditioning films, and the ambiguous fate of its protagonist. In holding a black mirror to the easy orthodoxies of every age, Burgess's masterpiece demonstrates that the dialectic of control is fought not only in state prisons but in the language we speak, the art we consume, and the stories we choose to tell about our own humanity.

References

Booker, M. K. (1996). *A practical introduction to literary theory and criticism*. Longman Publishers.

Böll, S. (2022). "This new Bogeyman of modern society": Anthony Burgess's A Clockwork Orange and the Figure of the Juvenile Delinquent in Postwar Britain. *European Journal of English Studies, 26*(1), 72–86.

Burgess, A. (1962). *A clockwork orange*. Penguin Books.

Cohen, S. (1972). *Folk Devils and Moral Panics: The Creation of the Mods and Rockers*. MacGibbon and Kee.

Cox, J. N., & Reynolds, L. J. (Eds.). (1993). *New historical literary study: Essays on reproducing texts, repressing history*. Princeton University Press.

Evans, R. O. (1971). Nadsat: The argot and its implications in Anthony Burgess' A clockwork orange. *Journal of Modern Literature*, *1*(3), 406–410.

Foucault, M. (1977). *Discipline and punish: The birth of the prison* (A. Sheridan, Trans.). Vintage Books.

Foucault, M. (1978). *The history of sexuality*, Vol. 1 (R. Hurley, Trans.). Pantheon Books.

Galbraith, J. K. (1958). *The affluent society*. Houghton Mifflin.

Geertz, C. (1973). *The interpretation of cultures: Selected essays*. Basic Books.

Geertz, C. (1980). *Negara: The theatre state in nineteenth-century Bali*. Princeton University Press.

Geertz, C. (1983). *Local knowledge: Further essays in interpretive anthropology*. Basic Books.

Gehrke, P. J. (2001). Deviant subjects in Foucault and "A clockwork orange": Congruent critiques of criminological constructions of subjectivity. *Critical Studies in Media Communication*, 18(3), 270–284. https://doi.org/10.1080/07393180128088

Greenblatt, S. (1980). *Renaissance self-fashioning: From More to Shakespeare*. University of Chicago Press.

Greenblatt, S. (1989). Towards a poetics of culture. In H. A. Veeser (Ed.), *The new historicism* (pp. 1–14). Routledge.

Greenblatt, S. (1991). *Learning to curse: Essays in early modern culture*. Routledge.

Hall, S. (1959). The politics of adolescence. *Universities and Left Review*, 6: 2.

Jackson, R. (2018). Technology of behavior: A clockwork orange and the philosophy of control. *Political Studies Review, 16*(3), 235-245.

Menking, L. (2012). Indeterminate mantras: Finding autonomy in *Slaughterhouse-Five* and *A Clockwork Orange*. *Second Line - An Undergraduate Journal of Literary Conversation*, 2012. http://www.secondlineblog.org/2012/06/indeterminate-mantras-finding-autonomy-in-slaughterhouse-five-and-a-clockwork-orange/

Montrose, L. (1986). Renaissance literary studies and the subject of history. *English Literary Renaissance*, *16*(1), 5–12.

Montrose, L. (1989). Professing the Renaissance: The poetics and politics of culture. In H. A. Veeser (Ed.), *The new historicism* (pp. 15–36). Routledge.

Montrose, L. (1992). New historicisms. In S. Greenblatt & G. Gunn (Eds.), *Redrawing the boundaries: The transformation of English and American literary studies* (pp. 392–418). Modern Language Association.

Newman, B. (1991). *A clockwork orange*: Burgess and behavioral interventions. *Behavior and Social Issues*, *1*(2), 61-70. https://doi.org/10.5210/bsi.v1i2.170

Skinner, B. F. (1953). *Science and human behavior*. Macmillan.

Skinner, B. F. (1971). *Beyond freedom and dignity*. Alfred A. Knopf.

Van Lier, M. (2018). Kubrick's Omission: The Twenty-First Chapter and the Critical Reception of *A Clockwork Orange*. *Adaptation, 11*(2), 198–215.

Veeser, H. A. (1989). Introduction. In H. A. Veeser (Ed.), *The new historicism* (pp. ix–xvi). Routledge.

Author's Biography

Pouyan Rezapour received his B.A., M.A., and Ph.D. degrees in English Language and Literature from Lorestan University, the University of Isfahan, and University of Tehran, respectively. In 2014, Rezapour joined the Department of English, Lorestan University, as a part-time Lecturer and from 2024 he has become a full-time Lecturer on a wide array of Literary studies both in B.A. and M.A. level. Since then, he has had over 10 publications on Comparative literary studies, Postcolonial reading of Literary text, sociopolitical studies, and English literature in domestic and international journals. In his initial year of Emplyment he has begun to supervise 5 MA students. Since 2024, he has also been the Director of Kuhdasht Higher Education Center.

Volume 14, Issue 1
Winter and Spring, 2026
pp. 283-286

Book review: Nassaji, H., *Applied Linguistics Research: A Comprehensive Guide to Methodology, Design, Analysis, and Evaluation*, Cambridge University Press, 2025, 824 pages, ISBN: 9781108672146 https://doi.org/10.1017/9781108672146

Elham Farahani*

*Assistant Professor of TEFL, Department of English Language and Literature, Faculty of Letters and Languages, Arak University, Arak, Iran
ORCID: 0009-0001-9068-755X; Email: e-farahani@araku.ac.ir*

Teaching research methodology in applied linguistics presents a familiar pedagogical challenge: the search for a resource that effectively bridges the theoretical foundations of inquiry with the practical demands of novice researchers. Too often, the subject is perceived as a challenging topic, with dense, voluminous texts overwhelming graduate students. It was from this dual standpoint, as an instructor seeking a user-friendly textbook and as a researcher desiring a clear reference, that I approached Hossein Nassaji's book on applied linguistics research. The text positions itself as that essential bridge, claiming to serve both experienced researchers and newcomers eager to begin their research journey. A thorough analysis of this book reveals that it largely fulfills this promise, offering a process-based approach that is particularly valuable for classroom instruction, albeit with certain areas ripe for refinement.

The book's twenty-five chapters are systematically organized into seven thematic parts, constructing a roadmap for the research process. Part I: Introduction to Applied Linguistics Research establishes a necessary foundation, addressing the nature of research, the scope of the field, and ethical principles. As in most books in this area, the first chapter focuses on the definition of research; compared with other available texts, however, this beginning delves into the world of research in a more engaging way. Chapter two introduces the nature, significance, and scope of applied linguistics. Chapter three discusses ethics in applied linguistics research. This thorough chapter is very informative; however, it would be better if the ethical issues of each part of the research process were introduced in the related chapter, rather than gathering all the points in a single chapter. For example, the ethics related to data collection could have been discussed in the chapter devoted to this stage of research, or plagiarism could be mentioned in the final chapter of the book that is related to writing the research report. More importantly, this chapter does not refer to the problems related to the use of AI.

Part II: The Research Process and Designing is where the text's pedagogical planning becomes distinctive. Chapter four introduces a valuable ten-step model of research, providing a clear, iterative scaffold. The model consists of four major parts: identification, preparation, implementation, and communication.

Chapter five presents a multidimensional framework for classifying research, with different types across six dimensions: paradigm, methodology, design, aim, applicability, and time frame. One issue with this chapter is that the design dimension is not clear; this category remains ambiguous, and all the examples of designs are from a quantitative approach. It would have been better if the author had provided examples of the term "design" from qualitative and mixed-methods approaches to avoid any misunderstanding. Based on my experience teaching this topic in research classes, most students struggle to distinguish between terms like approach, design, methodology, and methods due to varying definitions and uses across sources. Chapter six covers the difference between constitutive definitions and operational definitions, along with variables and scales in research. Chapter seven addresses the literature review and how to develop research questions. It defines the concept of the literature review and explains its aim and importance in the research process, outlines the steps involved in conducting an effective literature review, and distinguishes between various types of literature review, identifying their characteristics and purposes. Here, it distinguishes four main types of reviews: a traditional or narrative review, a systematic review, a meta-analysis, and a meta-synthesis. It also identifies the key features that define a high-quality literature review, discusses preliminary versus more comprehensive literature reviews, explores the literature review in qualitative and quantitative studies, examines different sources of research questions, explains how to develop clear, concise, and effective research questions, recognizes the differences between research questions for qualitative and quantitative research, and outlines the criteria for assessing good research questions. This chapter stands out because the issues related to conducting a literature review are presented cohesively and more effectively than in other sources. For most beginner researchers, conducting a literature review is very difficult.

Building on the foundational elements introduced in Part II, the book transitions into the specifics of quantitative research designs and data collection in Part III, focusing on the core aspects of conducting quantitative studies in applied linguistics through five chapters that guide readers through participant selection, research design, data collection, and ensuring research quality via reliability and validity assessments, all with an emphasis on practical techniques tailored to quantitative paradigms. Specifically, Chapter 8 explores various sampling strategies for selecting participants, ensuring representativeness and appropriateness for quantitative studies, while Chapter 9 discusses different types of quantitative research designs, including experimental, correlational, and survey-based approaches, along with techniques for implementing them effectively. Chapter 10 covers tools such as questionnaires, tests, and observational instruments, as well as techniques for gathering numerical data reliably. Meanwhile, Chapter 11 explains the concept of reliability, including types like test-retest and internal consistency, and methods for assessing and improving it, and Chapter 12 focuses on validity concepts, such as content, construct, and criterion validity, with strategies for evaluation to ensure research measures what it intends.

This structured approach is effectively mirrored in Part IV: Qualitative Research Designs, Data Collection, and Analysis, which parallels the previous quantitative focus but shifts to qualitative methods, offering in-depth treatments of analogous topics while incorporating data analysis, across five chapters that emphasize interpretive methods, strategies for data handling, and establishing rigor in qualitative research. In particular, Chapter 13 details purposeful and theoretical sampling techniques suited for in-depth qualitative inquiries, differing from quantitative random sampling, and Chapter 14 outlines qualitative design types like case studies, ethnography, and grounded theory, along with techniques for structuring flexible, exploratory research. Chapter 15 describes methods such as interviews, focus groups, observations, and document analysis, with strategies for ethical and effective data gathering, while Chapter 16 provides guidance on qualitative analysis techniques, including thematic analysis, content analysis, and coding strategies to interpret non-numerical data. Finally, Chapter 17 addresses concepts of trustworthiness in qualitative contexts, such as credibility, transferability, dependability, and confirmability, with methods to establish them, such as triangulation and member checking.

A particular strength emerges in Part V: Quantitative Research: Data Analysis and Interpretation. This six-chapter sequence functions as an exceptionally practical workshop manual, demystifying statistics from descriptive basics through to generalized linear models and effect size interpretation with uncommon clarity. It covers analyzing quantitative data with descriptive statistics, exploring key concepts in inferential statistics, commonly used inferential statistics tests in applied linguistics research (both parametric and nonparametric), advanced statistical models including GLMs, mixed models, and multivariate analysis, and effect sizes with types, calculations, and interpretation. This part of the book distinguishes it most from other research books in applied linguistics, since it provides readers with concrete examples and serves like a workshop teaching data analysis and SPSS in quantitative research.

The final sections provide essential closure: Part VI thoughtfully integrates paradigms in mixed methods research, and Part VII adeptly guides the dissemination of findings. This comprehensive, linear progression is a principal strength, rendering a complex process both manageable and learnable.

The book's standout achievement is its accessible, teacher-informed prose. It avoids becoming an abstract treatise, instead operating as a practical handbook. The text's pedagogical features such as succinct summaries, tables, and diagrams, are carefully crafted to encourage engagement with complex ideas. Several features offer immediate utility: abstracts and concrete examples for each design help students visualize end products; discussions of contemporary tools (e.g., transcription software) address real-time needs; and the quantitative analysis section is arguably the most instructively clear in any comparable textbook.

Nevertheless, a close reading from an instructor's perspective suggests avenues for enhancement in future editions. First, the early conceptual chapters would benefit from more balanced illustrative examples across paradigms to avoid unintentionally marginalizing qualitative approaches at a critical juncture. Second, the end-of-chapter "Discussion and Activity Questions" could be more strategically structured and applied. Placing discussion questions at the outset to prime reflection, followed by hands-on activities such as analyzing mini-datasets or evaluating sample literature excerpts, would strengthen pedagogical impact. Third, while a dedicated ethics chapter is valuable, weaving specific ethical considerations (e.g., digital data stewardship, AI use) into relevant process chapters might reinforce their practical import more effectively.

Furthermore, for a 2025 publication, the inclusion of more recent exemplars would better reflect evolving scholarly trends. More critically, the reference list does not adhere to the latest APA style edition, a notable oversight for a methods textbook destined to serve as a writing model. A simple yet powerful addition would be a curated list of exemplary published studies at the end of each design chapter, allowing students to immediately connect principles to practice.

Despite these critiques, Nassaji's contribution is significant and timely. It succeeds in demystifying methodological complexity, translating it into a coherent and teachable sequence.

I highly recommend Applied Linguistics Research as a core textbook for graduate methodology courses in applied linguistics and second language studies. It stands as one of the most accessible and comprehensive practical guides currently available and will also serve as a reliable desk reference for teacher-researchers. I intend to adopt it for my own courses and look forward to subsequent editions that, by incorporating more balanced examples and refined pedagogical tools, may well solidify its status as a definitive practical guide in our field.

Author's Biography

Elham Farahani received her BA in English Language and Literature from the University of Qom, and her MA and PhD degrees in TEFL from Arak University. She has served as a faculty member from 2018 to 2022 in the Department of English Language at Mofid University. In 2023, she joined Arak University, where she is currently an Assistant Professor. Her areas of interest include Corpus Linguistics, English for Academic Purposes, English for Research Publication Purposes, and (Critical) Discourse Analysis.

Volume 14, Issue 1
Winter and Spring, 2026
pp. 287-298

Journal of Applied Linguistics and Applied Literature: Dynamics and Advances

Persian Abstracts:

ارزیابی ارزیابان: مطالعه‌ای بر دیدگاه‌ها و شیوه‌های ارزیابی نگارش معلمان زبان انگلیسی به عنوان زبان خارجی سطح ۲ دانشگاه تگزاس و تأثیر آن بر نمرات نگارش دانش‌آموزان

زهرا زرگران (نویسنده مسئول)
دکترای آموزش زبان انگلیسی، دانشگاه فناوری و علوم کاربردی، شناص، عمان
Email: Zahra.zaragaran@utas.edu.om

محسن قربانپور
کارشناسی ارشد آموزش زبان انگلیسی، دانشگاه فناوری و علوم کاربردی، المصنعه، عمان
Email: Mohsen.ghorbanpoor@utas.edu.om

چکیده

این مطالعه به بررسی برداشت‌ها و شیوه‌های معلمان زبان انگلیسی به عنوان زبان خارجی سطح ۲ در استفاده از روبریک‌های استاندارد ارزیابی نگارش در دانشگاه فناوری و علوم کاربردی (UTAS) در عمان می‌پردازد. از یک مصاحبه نیمه‌ساختار یافته و پروتکل تفکر با صدای بلند برای بررسی نحوه تفسیر و اعمال معیارهای روبریک توسط معلمان هنگام نمره‌دهی به بخش نوشتاری تاسک ۲ استفاده شد. از تحلیل موضوعی برای تجزیه و تحلیل داده‌های کیفی استفاده شد و نقاط قوت و محدودیت‌های روبریک فعلی، مانند توصیف‌گرهای مبهم و عدم تطابق با انتظارات زبان‌آموز سطح پیش متوسط، آشکار شد. بر اساس بازخورد معلمان، یک روبریک اصلاح‌شده توسعه داده شد. برای ارزیابی تأثیر آن، از یک طرح درون‌گروهی استفاده شد و آزمون t با نمونه‌های جفت‌شده برای مقایسه نمرات نوشتاری دانش‌آموزان در مقایسه با دو نسخه روبریک انجام شد. نتایج نشان دهنده بهبود قابل توجه در نمرات کلی نوشتاری دانش‌آموزان به ویژه در زمینه‌های دستور زبان و واژگان بود. یافته‌ها بر ارزش طراحی روبریک آگاهانه توسط معلم، تطابق با اهداف آموزشی و تعدیل مداوم تأکید دارند. این مطالعه به افزایش پایایی ارزیابی و ارتباط آموزشی در ارزیابی نوشتاری زبان انگلیسی به عنوان زبان خارجی در مؤسسات آموزش عالی مانند UTAS در عمان کمک می‌کند.

واژه‌های کلیدی: روبریک نگارش، UTAS، عمان، آموزش نگارش، ارزیابی نگارش

اطلاعات مقاله

مقاله پژوهشی
تاریخ دریافت: شنبه، ۲۰ شهریور ۱۴۰۴
تاریخ تصویب: یکشنبه، ۱۲ بهمن ۱۴۰۴
تاریخ انتشار: چهارشنبه، ۱۲ فروردین ۱۴۰۵
تاریخ آنلاین: یکشنبه، ۱۲ بهمن ۱۴۰۴
شناسهٔ دیجیتال مقاله: https://doi.org/ 10.22049/jalda.2026.30990.1878

شاپای‌الکترونیکی: ۲۸۲۱-۰۲۰۴ شاپای‌چاپی: ۲۸۲۰-۸۹۸۶

اثر تمرکز بر فرم در رویکرد آموزش مبتنی بر فعالیت بر سلاست گفتاری و انگیزش زبان آموزان

سیمین عنبرشاهی (نویسنده مسئول)
استادیار آموزش زبان انگلیسی، گروه زبان انگلیسی، دانشکده ادبیات و علوم انسانی، دانشگاه شهید مدنی آذربایجان، تبریز، ایران
Email: anbarshahi@yahoo.com

امید یگانه
دانش آموخته کارشناسی ارشد آموزش زبان انگلیسی، دانشگاه شهید مدنی آذربایجان، تبریز، ایران
Email: omidyeganeh44@yahoo.com

چکیده

آموزش زبان مبتنی بر فعالیت (TBLT) نقش برجسته و چشمگیری در فرآیند آموزش زبان داشته و نظر بسیاری از محققان را به خود جذب کرده است (الیس، ۲۰۰۳؛ نونان، ۲۰۰۴؛ پرابهو، ۱۹۸۷؛ اسکهان، ۱۹۹۶). ضرورت داشتن سطح مقبولی از تسلط بر زبان انگلیسی زبان آموزان ایرانی نیز اخیراً به موضوعی بحث برانگیز تبدیل شده است. طبق استدلال محققان ارتباطات شفاهی (گفتاری) و تعامل نقشی حیاتی در یادگیری زبان ایفا می کند (الیس ۲۰۰۳ و نونان ۲۰۰۶). نونان (۱۹۹۲) پیشنهاد می‌کند که یادگیری زبان مبتنی بر فعالیت می‌تواند سطح و سرعت گفتمان دانش‌آموزان را افزایش داده، فضایی آرام بخش در کلاس ایجاد و ورودی قابل درک دانش‌آموزان را تقویت کند. زبان انگلیسی در ایران به عنوان یک زبان خارجی در نظر گرفته می‌شود که به این معنی است که در مواجهه با ورودی انگلیسی، محدودیت وجود دارد و در نتیجه، فراگیران در یادگیری و برقراری ارتباط از طریق زبان انگلیسی در جایی غیر از فضای کلاس با مشکل مواجه می‌شوند. و همانطور که تیل و گودرزی (۲۰۱۵) ارائه کردند، فرصت بالقوه ای برای فراگیران وجود دارد تا از فعالیت‌هایی که در کلاس به آن‌ها داده می شود، به نحو احسن و آموزنده استفاده کنند. بنابراین، تمرکز اصلی این مقاله بر روی تأثیر TBLT تمرکز بر فرم، بر انگیزه و روان زبان انگلیسی زبان‌آموز ایرانی است. طرح این پژوهش بصورت شبه آزمایشی بوده که در یک موسسه در تبریز، آذربایجان شرقی، ایران و بر روی گروه آزمایش و کنترل انجام شد و آن‌ها دوره ای ۱۲ جلسه ای را برای بررسی تأثیر این رویکرد گذراندند. در طول تحقیق، تسلط گفتاری آنها با معیارهای تسلط نرخ A و نرخ B بر اساس دستورالعمل های ارائه شده ی اندرسون و لینچ (۱۹۸۸) ارزیابی شد. به شرکت کنندگان در گروه آزمایش در سه فاز فرصت استفاده از زبان در شرایط واقعی داده می شد، در حالیکه گروه کنترل تمرینات تکراری و طوطی واری داده می شد. به منظور توضیح بیشتر در مورد موضوع انگیزش طرحی که در دورنی (۲۰۰۳) ارائه داده است، برای ارزیابی انگیزه تحصیلی شرکت کنندگان یک بار قبل از مطالعه و یک بار بعد از مطالعه استفاده شد. این نسخه از پرسشنامه شامل دو بخش عمده است: بخش اول شامل مواردی است که نگرش و انگیزه زبان آموزان را در مورد یادگیری زبان انگلیسی می سنجد و بخش دوم شامل سؤالاتی در مورد اطلاعات پیشینه زبان آموزان (تاگوچی و همکاران، ۲۰۰۹) است.

واژگان کلیدی: آموزش زبان مبتنی برفعالیت، روانی گفتار، انگیزش، تمرکز بر فرم

اطلاعات مقاله

مقالهٔ پژوهشی
تاریخ دریافت: شنبه، ۷ خرداد ۱۴۰۴
تاریخ تصویب: پنجشنبه، ۲۱ اسفند ۱۴۰۴
تاریخ انتشار: چهارشنبه، ۱۲ فروردین ۱۴۰۵
تاریخ آنلاین: پنجشنبه، ۲۱ اسفند ۱۴۰۴
شناسهٔ دیجیتال مقاله: https://doi.org/ 10.22049/jalda.2026.30662.1813

شاپای‌الکترونیکی: ۰۲۰۴-۲۸۲۱ شاپای‌چاپی: ۲۸۲۰-۸۹۸۶

تعامل هوش هیجانی و خودکارآمدی در پیش‌بینی تنظیم هیجان در فراگیران زبان انگلیسی در ایران

علی عرب مفرد (نویسنده مسئول)
استادیار زبانشناسی کاربردی، گروه زبان و ادبیات انگلیسی، دانشکده علوم انسانی و اجتماعی، دانشگاه گلستان، گرگان، ایران
E-mail: a.arabmofrad@gu.ac.ir

یگانه کاظمی
کارشناس ارشد آموزش زبان انگلیسی، واحد علوم و تحقیقات گلستان، دانشگاه آزاد اسلامی، گرگان، ایران
Email: teflkzy@gmail.com

میلاد علیزاده
دانشجوی دکتری آموزش زبان انگلیسی، دانشکده پردیس، دانشگاه ارومیه، ارومیه، ایران
Email: milad.alizadeh.d@gmail.com

عیسی نیازی
استادیار مدیریت بازرگانی، گروه مدیریت، دانشکده علوم انسانی و اجتماعی، دانشگاه گلستان، گلستان، ایران
Email: e.niazi@gu.ac.ir

چکیده

این پژوهش به بررسی تعامل بین هوش هیجانی (EI)، خودکارآمدی (SE) و تنظیم هیجان (ER) در میان ۲۵۶ زبان‌آموز ایرانی سطح متوسط انگلیسی با سن ۱۰ تا ۱۵ سال می‌پردازد تا مشخص کند آیا خودکارآمدی و هوش هیجانی می‌توانند تنظیم هیجان را پیش‌بینی کنند. برای تحلیل داده‌ها از مدل‌سازی معادلات ساختاری (SEM) استفاده شد. نتایج نشان داد که رابطه مثبت معناداری بین هوش هیجانی و تنظیم هیجان وجود دارد، که بیانگر آن است زبان‌آموزان با هوش هیجانی بالاتر بهتر قادر به تنظیم هیجانات خود هستند. با این حال، رابطه معناداری بین خودکارآمدی و تنظیم هیجان یافت نشد، که نشان می‌دهد خودکارآمدی ممکن است به‌طور مستقیم بر تنظیم هیجان در این زمینه تأثیرگذار نباشد. علاوه بر این، همبستگی منفی متوسطی بین خودکارآمدی و هوش هیجانی مشاهده شد که حاکی از شکاف احتمالی بین اعتمادبه‌نفس زبان‌آموزان در توانایی‌هایشان و آگاهی هیجانی آن‌هاست. این یافته‌ها بر اهمیت پرورش هوش هیجانی در یادگیری زبان تأکید می‌کنند، زیرا این عامل نقش حیاتی در مدیریت چالش‌های هیجانی یادگیری یک زبان جدید دارد.

واژگان کلیدی: هوش هیجانی، خودکارآمدی، تنظیم هیجان، زمینه آموزش زبان انگلیسی، مدل‌سازی معادلات ساختاری

اطلاعات مقاله

مقالهٔ پژوهشی
تاریخ دریافت: دوشنبه، ۲۶ خرداد ۱۴۰۴
تاریخ تصویب: یکشنبه، ۲۶ بهمن ۱۴۰۴
تاریخ انتشار: چهارشنبه، ۱۲ فروردین ۱۴۰۴
تاریخ آنلاین: یکشنبه، ۲۶ بهمن ۱۴۰۴
شناسهٔ دیجیتال مقاله: https://doi.org/ 10.22049/jalda.2026.30570.1795
شاپای‌الکترونیکی: ۲۸۲۱-۰۲۰۴ شاپای‌چاپی: ۲۸۲۰-۸۹۸۶

پژوهشی پیکره‌محور درمورد اصطلاحات در مقاله‌های پژوهشی رشته زبان‌شناسی کاربردی

الهام فراهانی (نویسنده مسئول)

استادیار آموزش زبان انگلیسی، گروه زبان وادبیات انگلیسی، دانشکده ادبیات و زبان ها، دانشگاه اراک، اراک، ایران

Email: e-farahani@araku.ac.ir

هوشنگ یزدانی

دانشیار آموزش زبان انگلیسی، گروه زبان وادبیات انگلیسی، دانشکده ادبیات و زبان ها، دانشگاه اراک، اراک، ایران

Email: h-yazdani@araku.ac.ir

چکیده

پژوهش های اخیر نشان می دهد اصطلاحات که اغلب، وجه غیررسمی، پررنگ و لعاب و سرگرم کننده زبان محسوب می شدند، به گفتمان رسمی هم راه یافته اند، اما تاکنون درباره ی ورود اصطلاحات به گونه های نوشتاری و رسمی، ازجمله مقاله های پژوهشی، تحقیقی صورت نگرفته است. با توجه به این خلا، پژوهش حاضر، پییکره ای زبانی متشکل از ۶۴۰ مقاله ی پژوهشی (۵۶۷۵۵۵۴ واژه) که به صورت تصادفی از معتبرترین نشریات پژوهشی در گستره ی زبانشناسی کاربردی انتخاب شده اند را مورد بررسی قرار داد. هدف اول این پژوهش ارائه ی چارچوبی مناسب برای تشخیص و تعریف عملیاتی اصطلاحات آکادمیک بود که تاکنون مورد غفلت قرارگرفته اند. هدف دوم، ایجاد فهرستی از رایج ترین اصطلاحات آکادمیک در مقاله های پژوهشی در گستره ی زبان شناسی کاربردی بود. این پژوهش همچنین به بررسی هرگونه تغییرات احتمالی در بسامد اصطلاحات در مقاله ها پژوهشی از سال ۱۹۸۰-۲۰۲۰ میلادی پرداخت. نتایج بررسی های پیکره محور نشان داد که اصطلاحات انگلیسی در مقاله ها ی پژوهشی مورد استفاده قرار می گیرند لذا باید بیشتر مورد بررسی و توجه قرار بگیرند. علاوه بر این، بررسی در زمانی اصطلاحات در این گونه ی رسمی و آکادمیک نشان داد که رابطه ای بین تغییر زمان و بسامد اصطلاحات وجود ندارد و استفاده از اصطلاحات بیشتز مربوط به سبک نوشتاری نویسندگان مقاله ها است. یافته های این پژوهش میتواند مورد استفاده ی استادان دانشگاهی درسهای نگارش انگلیسی، دانشجویان مقاطع تحصیلی کارشناسی ارشد و دکتری و تهیه کنندگان کتابهای آموزشی در حوزه ی آموزش اصطلاحات و نگارش انگلیسی قرار گیرد.

واژگان کلیدی: اصطلاحات، پیکره ی زبانی، مقاله های پژوهشی، زبانشناسی کاربردی

اطلاعات مقاله

مقالهٔ پژوهشی

تاریخ دریافت: یکشنبه، ۱۱ آبان ۱۴۰۴

تاریخ تصویب: یکشنبه، ۲۴ اسفند ۱۴۰۴

تاریخ انتشار: چهارشنبه، ۱۲ فروردین ۱۴۰۵

تاریخ آنلاین: یکشنبه، ۲۴ اسفند ۱۴۰۴

شناسهٔ دیجیتال مقاله: https://doi.org/ 10.22049/jalda.2026.31176.1916

شاپای‌الکترونیکی: ۲۸۲۱-۰۲۰۴ شاپای‌چاپی: ۲۸۲۰-۸۹۸۶

کاوشی در پذیرش چت‌جی‌پی‌تی توسط زبان‌آموزان عراقی زبان انگلیسی به‌عنوان زبان خارجی برای یادگیری زبان: پژوهشی بر اساس مدل پذیرش فناوری

احمد راوضان سلمان
پژوهشگر پسا دکتری آموزش زبان انگلیسی، دانشگاه رازی، کرمانشاه، ایران
Email: ahmed.rawdhan@uomus.edu.iq

سامان عبادی (نویسنده مسئول)
استاد تمام آموزش زبان انگلیسی، گروه آموزش زبان و زبانشناسی کاربردی، دانشکده ادبیات و علوم انسانی، دانشگاه رازی، کرمانشاه، ایران
E-mail: samanebadi@gmail.com

سجاد ولایتی
دانشجوی دکتری آموزش زبان انگلیسی، دانشگاه رازی، کرمانشاه، ایران
Email: velayati.sajad@gmail.com

چکیده
این پژوهش با بهره‌گیری از رویکرد آمیخته تبیینیِ دنباله‌دار، به بررسی متغیرهایی پرداخت که پذیرش و استفاده از چت‌جی‌پی‌تی برای یادگیری زبان انگلیسی را در میان ۲۲۰ دانشجوی عراقی تبیین می‌کنند. بر اساس مدل پذیرش فناوری(TAM) ، در مرحله نخست از پرسش‌نامه‌ای برای سنجش برداشت از سودمندی، برداشت از سهولت استفاده، نگرش، قصد رفتاری و استفاده واقعی، به‌همراه لذت ادراک‌شده، شرایط تسهیل‌کننده و پیچیدگی فناوری استفاده شد. در مرحله بعدی، مصاحبه‌های نیمه‌ساختاریافته انجام گرفت. نتایج تحلیل کمّی نشان داد که تجربه پیشین با فناوری، پیش‌بینی‌کننده‌ای معنادار برای قصد رفتاری و نیز استفاده واقعی از فناوری است؛ به بیان دیگر، آمادگی دیجیتال عاملی مهم در پذیرش به‌شمار می‌رود. در مقابل، جنسیت متغیری معنادار در پذیرش نبود. داده‌های کیفی الگویی روشن را آشکار ساختند؛ بدین معنا که زبان‌آموزان انگلیسی به‌عنوان زبان خارجی، چت‌جی‌پی‌تی را ابزاری می‌دانند که به بهبود مهارت زبانی، افزایش انگیزش و ارائه پیشنهادهای شخصی‌سازی‌شده در یادگیری یاری می‌رساند. با وجود مزایایی همچون امکان خودارزیابی، نگرانی‌های قابل‌توجهی نیز نسبت به فقدان بازخورد تعاملی پویا، خطر وابستگی بیش از اندازه و ملاحظات اخلاقی وجود دارد. یافته‌ها نشان داد که زبان‌آموزان عراقی، چت‌جی‌پی‌تی را ابزاری سودمند تلقی می‌کنند؛ بااین‌حال، برای تضمین کاربرد مؤثر و منصفانه آن، لازم است چارچوب‌های آموزشی طراحی شود تا سواد دیجیتال را حمایت کرده و دستورالعمل‌های روشنی برای کنترل چالش‌های عملی و اخلاقی فراهم آورد.

واژگان کلیدی: چت‌جی‌پی‌تی، جنسیت، زبان‌آموزان عراقی، رویکرد آمیخته، مدل پذیرش فناوری(TAM)

اطلاعات مقاله
مقالهٔ پژوهشی
تاریخ دریافت: یکشنبه، ۲۵ آبان ۱۴۰۴
تاریخ تصویب: یکشنبه، ۲۶ بهمن ۱۴۰۴
تاریخ انتشار: چهارشنبه، ۱۲ فروردین ۱۴۰۵
تاریخ آنلاین: یکشنبه، ۲۶ بهمن ۱۴۰۴
شناسهٔ دیجیتال مقاله: https://doi.org/ 10.22049/jalda.2026.31224.1923
شاپای‌الکترونیکی: ۲۸۲۱-۰۲۰۴ شاپای‌چاپی: ۲۸۲۰-۸۹۸۶

ارزیابی فعالیت‌های شنیداری در کتاب‌های آمادگی آزمون آیلتس

ساسان بالغی زاده (نویسنده مسئول)
استاد زبانشناسی کاربردی، گروه زبان و ادبیات انگلیسی، دانشکده ادبیات و علوم انسانی،
دانشگاه شهید بهشتی، تهران، ایران
Email: sasanbaleghizadeh@yahoo.com

مهران موسوی
دانشجوی دوره دکترای آموزش زبان انگلیسی، دانشگاه شهید بهشتی، تهران، ایران
E-mail: mehran.mousaviii@gmail.com

چکیده
اگرچه منابع آمادگی آزمون آیلتس غالباً مدعی ارتقای مهارت‌های ارتباطی زبان‌آموزان هستند، اما نگرانی فزاینده‌ای وجود دارد که این منابع بیش از آنکه به پرورش مهارت‌های شنیداری واقعی و کاربردی بپردازند، بر راهبردهای آزمون‌محور تمرکز دارند. پژوهش حاضر با هدف بررسی جهت گیری ارتباطی فعالیت‌های شنیداری در کتاب‌های آمادگی آزمون آیلتس که به طور گسترده در ایران مورد استفاده قرار می‌گیرند انجام شده است؛ بافتی که در آن، زبان‌آموزان به دلیل محدودیت دسترسی و مواجهه با زبان انگلیسی در محیط های طبیعی، به شدت متکی بر منابع آموزشی هستند. در این مطالعه، با بهره‌گیری از گونه‌شناسی فعالیت‌های شنیداری گوه (۲۰۰۱)، مجموعه‌ای گسترده از فعالیت‌های شنیداری در شش کتاب بین‌المللی آمادگی آیلتس مورد تحلیل قرار گرفت. فعالیت‌ها بر اساس ساختار مشارکتی و هدف ارتباطی در چهار گروه طبقه‌بندی شدند: تراکنشی یک طرفه، تعاملی یک طرفه، تراکنشی دوطرفه و تعاملی دوطرفه. یافته‌ها حاکی از وجود عدم توازن آشکار در طراحی این فعالیت‌ها بود؛ به‌گونه‌ای که بخش عمده‌ای از آن‌ها در دسته‌ی تراکنشی یک طرفه قرار داشته و بیشتر بر درک منفعلانه اطلاعات متمرکز بودند و فرصت‌های محدودی برای تعامل یا تفسیرپراگماتیک فراهم می‌کردند. هرچند برخی کتاب‌های مورد بررسی در سطوح زبانی بالاتر نشانه‌هایی از حرکت تدریجی به سوی فعالیت‌های ارتباطی‌تر را نشان می‌دهند، برخی دیگر همچنان بر رویکردی صرفاً آزمون محور استوارند. . این ناهمگونی نشان‌دهنده حمایت ناکافی از رشد مهارت‌های شنیداری کاربردی در موقعیت‌های واقعی است. در بافت آموزش زبان انگلیسی به‌عنوان زبان خارجی در ایران، چنین تمرکز یک جانبه‌ای بر فعالیت‌های غیرتعاملی می‌تواند به تضعیف توانایی‌های کاربردی و راهبردی زبان‌آموزان منجر شده و آن‌ها را به اندازه کافی برای ارتباطات آکادمیک و حرفه‌ای آماده نسازد. این پژوهش بر نیاز مبرم به طراحی منابع آموزشی و راهبردهای تدریس ارتباط محورتر تأکید دارد که با اهداف برنامه درسی ملی هم راستا بوده و مهارت‌های شنیداری زبان‌آموزان را برای کاربردهای واقعی تقویت کند.

واژگان کلیدی: منابع آمادگی آیلتس، فعالیت‌های شنیداری واقعی، مهارت‌های ارتباطی، اثر بازخوردی آزمون

اطلاعات مقاله

مقالهٔ پژوهشی
تاریخ دریافت: یکشنبه، ۶ مهر ۱۴۰۴
تاریخ تصویب: دوشنبه، ۲۵ اسفند ۱۴۰۴
تاریخ انتشار: چهارشنبه، ۱۲ فروردین ۱۴۰۴
تاریخ آنلاین: دوشنبه، ۲۵ اسفند ۱۴۰۴
شناسهٔ دیجیتال مقاله: https://doi.org/ 10.22049/jalda.2026.31055.1890
شاپای‌الکترونیکی: ۲۸۲۱-۰۲۰۴ شاپای‌چاپی: ۲۸۲۰-۸۹۸۶

خواندن با پشتیبانی: تأثیر مدل‌های حمایتی بر خودِ تحصیلی، تاب‌آوری و خودتنظیمی زبان‌آموزان زبان انگلیسی به عنوان زبان خارجی

کمیل جعفری سنکسرک

دانشجوی دکتری آموزش زبان انگلیسی، واحد قائمشهر، دانشگاه آزاد اسلامی، قائمشهر، ایران

komeil.jafari5227@iau.ac.ir

مهرشاد احمدیان (نویسنده مسئول)

استادیار آموزش زبان انگلیسی، گروه زبانهای خارجی، واحد قائمشهر، دانشگاه آزاد اسلامی، قائمشهر، ایران

Email: mehrshad.ahmadian@iau.ac.ir

ابراهیم فخری علمداری

استادیار آموزش زبان انگلیسی، گروه زبانهای خارجی، واحد قائمشهر، دانشگاه آزاد اسلامی، قائمشهر، ایران

Email: ebrahim.fakhri@qaemiau.ac.ir

چکیده

این مطالعه ترکیبی، تأثیر سه مدل آموزشی درک مطلب آکادمیک حمایتی - حمایت مبتنی بر کلاس درس با رویکرد اجتماعی-فرهنگی جویس-گیبونز (۲۰۱۷)، چارچوب حمایت مشروط ون د پل و همکاران (۲۰۱۰) و رویکرد حمایت کویل (۲۰۱۵) در زمینهٔ پیشرفت خواندن و ویژگی‌های انتقادی فراگیرنده – را بر پیشرفت خواندن، انگیزه، تاب‌آوری و خودتنظیمی دانشجویان کارشناسی در دانشگاه آزاد اسلامی مازندران بررسی نمود. در مجموع ۱۲۰ دانشجوی فارسی‌زبان، ۱۹ تا ۲۵ ساله، از رشته‌های تحصیلی مختلف، به طور تصادفی به سه گروه آزمایشی اختصاص داده شدند. ابزارهای استاندارد شامل آزمون تعیین سطح آکسفورد، پرسشنامه خودانگیزشی زبان دوم، پرسشنامه تاب‌آوری، پرسشنامه خودتنظیمی و آزمون درک مطلب قبل و بعد از مداخله اجرا شد. طی ۸ تا ۱۰ هفته، هر گروه آموزشی حمایت متناسب با چارچوب مدل اختصاص داده شده خود را دریافت کرد. تحلیل واریانس چندگانه (MANOVA) بهبودهای قابل‌توجهی را در تمام حوزه‌های اندازه‌گیری شده برای هریک ار سه مدل آموزشی در مقایسه با گروه شاهد نشان داد، که در آن ارقام مربوط به بزرگی تأثیر، نتایج قابل توجهی را در هر دو حوزهٔ شناختی و عاطفی زبان‌آموز تأیید می‌کند. هر یک از این مدل‌ها، اگرچه در مبانی آموزشی خود متمایز هستند، اما تعهد مشترکی به ارائه حمایت ساختاریافته، پاسخگو و تدریجی دارند که گذار زبان‌آموزان از عملکرد همراه با کمک به عملکرد مستقل را تسهیل می‌کند. داده‌های کیفی حاصل از مصاحبه‌های نیمه‌ساختاریافته، حاکی از برداشت‌های مثبت زبان‌آموزان از آموزش حمایتی و نقش آن در افزایش مشارکت تحصیلی و عاطفی آنها داشت. یافته‌ها بر ارزش آموزشی استراتژی‌های حمایتی خواندن در تقویت رشد جامع زبان‌آموز تأکید می‌کنند و دستورالعمل هایی را برای ادغام رویکردهای حمایتی منطبق با فرهنگ در محیط‌های یادگیری زبان انگلیسی به عنوان زبان خارجی ارائه می دهند.

واژگان کلیدی: پیشرفت تحصیلی، انگیزش، تاب‌آوری، آموزش حمایتی، خودتنظیمی

اطلاعات مقاله

مقالهٔ پژوهشی

تاریخ دریافت: سه شنبه، ۱۴ مرداد ۱۴۰۴

تاریخ تصویب: پنجشنبه، ۱۶ بهمن ۱۴۰۴

تاریخ انتشار: چهارشنبه، ۱۲ فروردین ۱۴۰۵

تاریخ آنلاین: پنجشنبه، ۱۶ بهمن ۱۴۰۴

شناسهٔ دیجیتال مقاله: https://doi.org/1 10.22049/jalda.2026.30880.1850

شاپای‌الکترونیکی: ۲۸۲۱-۰۲۰۴ شاپای‌چاپی: ۲۸۲۰-۸۹۸۶

پویایی‌های اضطراب فنآوری در مدرسان زبان انگلیسی به عنوان زبان خارجی: مدل‌سازی معادلات ساختاری مرتبه اول و دوم بر اساس خودکارآمدی رایانه‌ای، دانش محتوایی تربیتی فناورانه، فرسودگی شغلی، تاب‌آوری تحصیلی، حمایت و جمعیت‌شناسی

هدی دیوسر (نویسنده مسئول)

دانشیار آموزش زبان انگلیسی، گروه زبان و ادبیات انگلیسی، دانشکده ادبیات و علوم انسانی، دانشگاه گیلان، رشت، ایران

E-mail: hodadivsar@guilan.ac.ir

چکیده

با وجود نقش امیدوارکننده فناوری در آموزش، اضطراب فناوری (تکنواسترس) معلمان مانعی جدی برای ادغام فناوری در فرایند تدریس به شمار می‌آید. این پژوهش با هدف آزمون یک مدل ساختاری از تکنواسترس معلمان، دانش محتوایی تربیتی فناورانه (TPACK)، خودکارآمدی رایانه‌ای، حمایت محیطی، تاب‌آوری تحصیلی و فرسودگی شغلی انجام شد. برای دستیابی به این هدف، ۲۳۲ مدرس زبان انگلیسی به عنوان زبان خارجی «مقیاس خودکارآمدی رایانه‌ای»، «پرسشنامه تعیین سطح تکنواسترس معلمان»، «پرسشنامه TPACK»، «پرسشنامه حمایت کلی»، «مقیاس تاب‌آوری کانر–دیویسون» و «فهرست فرسودگی شغلی مزلک (۲۰۱۶)» را تکمیل کردند. تحلیل عاملی اکتشافی و تأییدی روی تمامی پرسشنامه‌ها انجام شد و شاخص‌های مطلوب روایی و پایایی را تأیید کرد که از زیرعامل‌های موجود در سازه‌ها حمایت نمود. مدل‌یابی معادلات ساختاری (SEM) نشان داد که: الف) TPACK و تاب‌آوری معلمان زبان انگلیسی تکنواسترس آن‌ها را به‌طور منفی پیش‌بینی می‌کنند، در حالی‌که فرسودگی شغلی تکنواسترس را به‌طور مثبت پیش‌بینی می‌کند؛ ب) حمایت اداری و خودکارآمدی رایانه‌ای از طریق نقش میانجیِ TPACK و تاب‌آوری، تکنواسترس معلمان زبان انگلیسی را به صورت منفی پیش‌بینی می‌کنند؛ ج) فرسودگی شغلی معلمان به‌عنوان یک میانجی میان حمایت اداری و تکنواسترس و همچنین میان خودکارآمدی رایانه‌ای و تکنواسترس عمل می‌کند، به‌طوری‌که با متغیرهای برون‌زاد نهفته رابطه‌ای منفی داشته و تأثیر پیش‌بینانه‌ی معناداری بر متغیر درون‌زاد نهفته دارد؛ د) افزون بر این، TPACK و تاب‌آوری به‌طور غیرمستقیم بر فرسودگی شغلی معلمان اثر گذاشتند و نشان دادند که سطوح بالاتر TPACK و تاب‌آوری با کاهش فرسودگی شغلی و تکنواسترس همراه است؛ ه) تکنواسترس با سن رابطه مثبت داشت، اما رابطه آن با جنسیت معنادار نبود. نتایج این تحلیل به ارائه توصیه‌های آموزشی برای مدرسان زبان انگلیسی، مربیان و سیاست‌گذاران انجامید.

واژگان کلیدی: اضطراب فناوری (تکنواسترس)، دانش محتوایی تربیتی فناورانه، حمایت محیطی، خودکارآمدی رایانه ای، متغیرهای جمعیت‌شناختی

اطلاعات مقاله

مقالهٔ پژوهشی

تاریخ دریافت: دوشنبه، ۵ آبان ۱۴۰۴

تاریخ تصویب: یکشنبه، ۱۰ اسفند ۱۴۰۴

تاریخ انتشار: چهارشنبه، ۱۲ فروردین ۱۴۰۵

تاریخ آنلاین: یکشنبه، ۱۰ اسفند ۱۴۰۴

شناسهٔ دیجیتال مقاله: https://doi.org/ 10.22049/jalda.2026.31100.1897

شاپای‌الکترونیکی: ۲۸۲۱-۰۲۰۴ شاپای‌چاپی: ۲۸۲۰-۸۹۸۶

تأثیر یادگیری ترکیبی و مجازی بر تمایل به برقراری ارتباط و مهارت گفتاری در زبان‌آموزان عراقی زبان انگلیسی به‌عنوان زبان خارجی: مطالعه ای آمیخته

سامی ریحان
دانشجوی دکتری آموزش زبان انگلیسی، دانشگاه اصفهان، اصفهان، ایران
Email: raihansami2018@gmail.com

سعید کتابی (نویسنده مسئول)
دانشیار زبانشناسی کاربردی، گروه زبان و ادبیات انگلیسی، دانشکده زبانهای خارجی، دانشگاه اصفهان، اصفهان، ایران
Email: ketabi@fgn.ui.ac.ir

ضرغام قبانچی
دانشیار زبانشناسی کاربردی، گروه زبان و ادبیات انگلیسی، دانشکده ادبیات و علوم انسانی، دانشگاه فردوسی مشهد، مشهد، ایران
Email: ghabanchi@um.ac.ir

چکیده

این پژوهش به بررسی تأثیر یادگیری ترکیبی (Blended Learning) و یادگیری مجازی (Virtual Learning) بر تمایل به برقراری ارتباط (Willingness to Communicate) و مهارت گفتاری زبان‌آموزان عراقیِ زبان انگلیسی به‌عنوان زبان خارجی پرداخت. در چارچوب یک رویکرد آمیخته، ۱۲۰ دانشجو از دانشگاه کربلا در این مطالعه مشارکت داشتند که به‌صورت تصادفی در سه گروه آموزش حضوری، مجازی و ترکیبی جای‌دهی شدند. داده‌های کمی از طریق پیش‌آزمون و پس‌آزمون گردآوری شد، در حالی‌که داده‌های کیفی از طریق مصاحبه‌های نیمه‌ساختاریافته با مدرسان زبان انگلیسی به‌دست آمد. یافته‌ها نشان داد که هر دو شیوه یادگیری ترکیبی و مجازی به‌طور معناداری تمایل به برقراری ارتباط زبان‌آموزان را افزایش دادند، بدون آن‌که تفاوت معناداری بین این دو روش مشاهده شود. با این حال، یادگیری ترکیبی در بهبود مهارت گفتاری نسبت به یادگیری مجازی اثربخشی بیشتری نشان داد. مدرسان عراقی زبان انگلیسی به‌طور کلی دیدگاه مثبتی نسبت به هر دو روش تدریس ابراز کردند، هرچند تمایل بیشتری به یادگیری ترکیبی داشتند. در مجموع، نتایج پژوهش حاکی از آن است که هر دو رویکرد مزایای قابل‌توجهی برای آموزش زبان انگلیسی به‌عنوان زبان خارجی فراهم می‌کنند، اما یادگیری ترکیبی به‌دلیل تلفیق مزایای آموزش حضوری و مجازی، به‌عنوان روشی اثربخش‌تر برجسته می‌شود. این یافته‌ها می‌تواند در طراحی راهبردهای مؤثر آموزش زبان انگلیسی با بهره‌گیری از شیوه‌های متنوع یادگیری مورد استفاده قرار گیرد.

واژگان کلیدی: آموزش زبان انگلیسی به‌عنوان زبان خارجی، تمایل به برقراری ارتباط، زبان‌آموزان عراقیِ زبان انگلیسی، مهارت گفتاری، یادگیری ترکیبی، یادگیری مجازی

اطلاعات مقاله

مقالهٔ پژوهشی
تاریخ دریافت: جمعه، ۱۴ شهریور ۱۴۰۴
تاریخ تصویب: جمعه، ۲۸ آذر ۱۴۰۴
تاریخ انتشار: چهارشنبه، ۱۲ فروردین ۱۴۰۵
تاریخ آنلاین: جمعه، ۲۸ آذر ۱۴۰۴
شناسهٔ دیجیتال مقاله: https://doi.org/ 10.22049/jalda.2026.30974.1875
شاپای‌الکترونیکی: ۲۸۲۱-۰۲۰۴ شاپای‌چاپی: ۲۸۲۰-۸۹۸۶

بررسی رابطه علی بین روش های آموزشی معلمان، جهت گیری رشد و توسعه واژگان با مدل سازی معادلات ساختاری

محدثه هاشمی اقدم
دانشجوی دکتری آموزش زبان انگلیسی، دانشگاه ارومیه ، ارومیه، ایران
Email: mohashemi076@gmail.com

سیما مدیرخامنه
دانشیار آموزش زبان انگلیسی، گروه زبان و ادبیات انگلیسی، دانشکده ادبیات و علوم انسانی، دانشگاه ارومیه، ارومیه، ایران
E-mail: sima.modirkhamene843@gmail.com

پرویز علوی نیا
دانشیار آموزش زبان انگلیسی، گروه زبان و ادبیات انگلیسی، دانشکده ادبیات و علوم انسانی، دانشگاه ارومیه، ارومیه، ایران
Email: pevinia2006@gmail.com

چکیده
این پژوهش به بررسی تعامل میان شیوه‌های تدریس (از جمله ارتباط‌مندی، بازخورد/پیش‌خوراند، و سازمان‌دهی و شفافیت) و پیشرفت واژگانی دانش‌آموزان پرداخت. علاوه بر این، با توجه به نقش میانجی‌گر گرایش به رشد (GO) در این تعامل، نحوۀ اثرگذاری شیوه‌های تدریس(TP) بر گرایش به رشد نیز مورد تحلیل قرار گرفت. نمونه ی آماری متشکل از ۱۱۰۵ دانش‌آموز پایۀ دهم دبیرستان در این مطالعه مشارکت داشتند و پرسشنامه‌هایی را تکمیل کردند که شامل نسخه‌های بومی‌سازی‌شده مقیاس‌های GO و TP و نیز یک آزمون واژگان بود. داده‌ها با استفاده از مدل‌یابی معادلات ساختاری حداقل مربعات جزیی تحلیل شد که در نهایت به برازش بهینۀ مدل کلی پژوهش انجامید. یافته‌ها نشان داد برداشت دانش‌آموزان از TP به‌طور معناداری با GO و پیشرفت تحصیلی آنان مرتبط است. همچنین مشخص شد که GO با دستاوردهای تحصیلی دانش‌آموزان تعامل دارد. افزون بر این، نتایج بیانگر آن بود که GO نقش میانجی در ارتباط علّی میان ادراک دانش‌آموزان از TP و پیشرفت تحصیلی آنان ایفا می‌کند. این نتایج نقش اساسی پرورش GO در میان دانش‌آموزان را برجسته ساخته و بر تأثیر شیوه‌های تدریس اثربخش در شکل‌دهی ذهنیت، رفتارهای هدف‌گذاری رشدی و موفقیت تحصیلی آنان تأکید می‌کند. یافته‌ها همچنین از این دیدگاه حمایت می‌کنند که مداخلات هدفمند برای تقویت GO می‌توانند به بهبود معنادار دستاوردهای آموزشی منجر شوند.

واژگان کلیدی: جهت گیری رشد، توسعه واژگان، شیوه تدریس، انگلیسی به عنوان زبان خارجی

اطلاعات مقاله
مقالۀ پژوهشی
تاریخ دریافت: شنبه، ۱ شهریور ۱۴۰۴
تاریخ تصویب: چهارشنبه، ۱۲ فروردین ۱۴۰۵
تاریخ انتشار: چهارشنبه، ۱۲ فروردین ۱۴۰۵
تاریخ آنلاین: چهارشنبه، ۱۲ فروردین ۱۴۰۵

شناسۀ دیجیتال مقاله: https://doi.org/ 10.22049/jalda.2026.30934.1865

شاپای‌الکترونیکی: ۰۲۰۴-۲۸۲۱ شاپای‌چاپی: ۲۸۲۰-۸۹۸۶

JALDA Volume 14, Issue 1, Winter and Spring, 2026, pp. 287-298

تاب‌آوری در دل ویرانی: خوانشی بوم‌شناختی-پسااستعماری از «گل‌برگ‌های خون» اثر نگوگی وا تیونگو

ریحانه گودرزی
کارشناس ارشد ادبیات انگلیسی، دانشگاه تهران، تهران ، ایران
Email: goudarzi.reyhaneh@ut.ac.ir

حسین نظری (نویسنده مسئول)
استادیار زبان و ادبیات انگلیسی، گروه زبان و ادبیات انگلیسی، دانشکده زبان‌ها و ادبیات خارجی، دانشگاه تهران، تهران، ایران
E-mail: nazarih@ut.ac.ir

چکیده

این پژوهش به ارائه خوانشی بوم‌شناختی–پسااستعماری از رمان «گل‌برگ‌های خون» (Petals of Blood) اثر نگوگی وا تیونگو می‌پردازد و آن را در بستر بحران‌های اجتماعی–اقتصادی کنیا در دوران پس از استعمار، که تحت تأثیر میراث‌های استعمار و تداوم تاثیرات نواستعماری تشدید شده‌اند، بررسی می‌نماید. تحلیل حاضر به واکاوی بازنمایی تاب‌آوری و تروما در بستر آشفتگی‌های زیست‌محیطی و اجتماعی–اقتصادی در این رمان می‌پردازد. استدلال می‌شود که رمان خُردجهانی از چشم‌انداز اجتماعی–سیاسی کنیا ارائه می‌نماید و بر تروماهای ناشی از بهره‌کشی (نوا)استعماری تأکید می‌ورزد. شخصیت‌های محوری رمان تجسم این تروماهای جمعی‌اند، زیرا با سرکوب، خشونت، و خیانت مواجه می‌شوند. زوال زیست‌محیطی ایلموروگ نمونه‌ای از پیامدهای شیوه‌های کشاورزی استعماری و سوءمدیریت رایج در دوران پس از استعمار است. این پژوهش نشان می‌دهد که دانش بومی و کنش‌های اکولوژیک، در تقابل آشکار با رویکردهای سودمحور، مسیرهایی برای التیام و تاب‌آوری فراهم می‌آورند. طبیعت در رمان نماد امید و مقاومت است، در حالی که بن‌مایه «خون» هم‌زمان نمایانگر خشونت و همبستگی جمعی است. در نهایت، رمان برهم‌کنش و هم‌زمانیِ ویرانی و باززایی را در کنیای پس از استعمار بازمی‌نمایاند و بر پیوند ناگسستنی عدالت زیست‌محیطی و عدالت اجتماعی تأکید می‌نماید؛ همچنین بر ضرورت بازیابی دانش بومی و کنش‌های پایدار به‌منزله شرط آشتی و تاب‌آوری پای می‌فشارد.

واژگان کلیدی: نقد بوم شناختی، نقد پسااستعماری، مقاومت، تخریب محیط زیست، گل‌برگ‌های خون، کنیای پسااستعماری

اطلاعات مقاله
مقالهٔ پژوهشی
تاریخ دریافت: جمعه، ۵ دی ۱۴۰۴
تاریخ تصویب: چهارشنبه، ۲۷ اسفند ۱۴۰۴
تاریخ انتشار: چهارشنبه، ۱۲ فروردین ۱۴۰۵
تاریخ آنلاین: چهارشنبه، ۲۷ اسفند ۱۴۰۴
شناسهٔ دیجیتال مقاله: https://doi.org/ 10.22049/jalda.2026.31347.1944
شاپای‌الکترونیکی: ۰۲۰۴-۲۸۲۱ شاپای‌چاپی: ۲۸۲۰-۸۹۸۶

دیالکتیک کنترل: کاوشی نو-تاریخ‌گرایانه در رمان «پرتقال کوکی» اثر آنتونی برجس

پویان رضاپور (نویسنده مسئول)
استادیار زبان و ادبیات انگلیسی، گروه زبان انگلیسی، دانشکده ادبیات و علوم انسانی، دانشگاه لرستان، خرم‌آباد، ایران
Email: rezapour.p@lu.ac.ir

چکیده
تاریخ‌گرایی نوین که برای به چالش کشیدن نقد تاریخی و فرمالیستی ظهور کرده است، خواستار آن است که آثار ادبی در چارچوب بافت‌های پیچیده اجتماعی-تاریخی خود تحلیل شوند و مفهوم متن زیبایی‌شناختی مستقل را رد می‌کند. این مقاله با استفاده از این دیدگاه میان‌رشته‌ای، رمان جنجالی «پرتقال کوکی» سال ۱۹۶۲ آنتونی برجس را بررسی می‌کند. این مقاله بر آن است که دیدگاه دیستوپیایی رمان از فرهنگ جوانان معتاد به خشونت افراطی و دولتی که از شرطی‌سازی رفتاری سرکوبگرانه استفاده را به تصویر بکشد. این نکته نه صرفاً به عنوان یک خیال‌پردازی نظری، بلکه به عنوان یک پرسش انتقادی از عمیق‌ترین اضطراب‌های بریتانیای پس از جنگ عمل می‌کند. این تحلیل با به چالش کشیدن اثر برجس در ورطه های جنگ سرد، سرمایه‌داری مصرفی رو به رشد، بحران در اقتدار سیاسی و نفوذ فزاینده رفتارگرایی اسکینری، نشان می‌دهد که چگونه رمان تنش‌های فرهنگی فراگیر آن دوره را تجزیه و تحلیل می‌کند. در نهایت، این مطالعه ادعا می‌کند که «پرتقال کوکی» یک سند مهم از دوره خود است، یک عمل نمادین که گفتمان‌های بسیار واقعی کنترل اجتماعی، وحشت اخلاقی و تضاد ایدئولوژیکی را که اوایل دهه ۱۹۶۰ را تعریف می‌کردند، جذب و به صورت هنری متحول می‌کند.

واژگان کلیدی: پرتقال کوکی، آنتونی برجس ، زمینه، تاریخ‌گرایی نوین

اطلاعات مقاله
مقالهٔ پژوهشی
تاریخ دریافت: پنجشنبه، ۲۰ شهریور ۱۴۰۴
تاریخ تصویب: یکشنبه، ۱۲ بهمن ۱۴۰۴
تاریخ انتشار: چهارشنبه، ۱۲ فروردین ۱۴۰۵
تاریخ آنلاین: یکشنبه، ۱۲ بهمن ۱۴۰۴
شناسهٔ دیجیتال مقاله: https://doi.org/ 10.22049/jalda.2026.30990.1878

شاپای‌الکترونیکی: ۲۸۲۱-۰۲۰۴ شاپای‌چاپی: ۲۸۲۰-۸۹۸۶

دانشگاه شهید مدنی آذربایجان

دو فصلنامهٔ *زبان‌شناسی کاربردی ادبیات کاربردی: پویش‌ها و پیشرفت‌ها*

صاحب امتیاز:
دانشگاه شهید مدنی آذربایجان

مدیر مسئول:
دکتر داود امینی

سردبیر:
دکتر کریم صادقی

مدیر داخلی:
دکتر رضا یل شرزه

نشانی:
کیلومتر ۳۵، جاده تبریز- مراغه، دانشگاه شهید مدنی آذربایجان، دانشکده ادبیّات و علوم انسانی، دفتر دو فصلنامه.

تلفاکس:
۰۴۱-۳۴۳۲۷۵۵۹

کدپستی:
۵۳۷۵۱۷۱۳۷۹

آدرس سایت نشریه:
URL: http://jalda.azaruniv.ac.ir

شمارهٔ انتشار:
سال چهاردهم، شماره اول

تاریخ انتشار: چهارشنبه ۱۲ فروردین ۱۴۰۵

بها: ۵۰۰۰۰۰ ریال

شمارگان: ۱۰۰ جلد

این دوفصلنامه با مجوز شمارهٔ ثبت ۹۱/۳۴۷۱۵ وزارت فرهنگ و ارشاد اسلامی چاپ و منتشر می‌شود.

شاپای‌الکترونیکی: ۲۸۲۱-۰۲۰۴
شاپای‌چاپی: ۲۸۲۰-۸۹۸۶

اعضای هیأت تحریریه بین المللی

دانشگاه شهید مدنی آذربایجان

دو فصلنامهٔ *زبان‌شناسی کاربردی ادبیات کاربردی: پویش‌ها و پیشرفت‌ها*

صاحب امتیاز:
دانشگاه شهید مدنی آذربایجان

مدیر مسئول:
دکتر داود امینی

سردبیر:
دکتر کریم صادقی

مدیر داخلی:
دکتر رضا یل شرزه

نشانی:
کیلومتر ۳۵، جاده تبریز- مراغه، دانشگاه شهید مدنی آذربایجان، دانشکده ادبیّات و علوم انسانی، دفتر دو فصلنامه.

تلفاکس:
۰۴۱-۳۴۳۲۷۵۵۹

کدپستی:
۵۳۷۵۱۷۱۳۷۹

آدرس سایت نشریه:
URL: http://jalda.azaruniv.ac.ir

شمارهٔ انتشار:
سال چهاردهم، شماره اول

تاریخ انتشار: چهارشنبه ۱۲ فروردین ۱۴۰۵

بها : ۵۰۰۰۰۰ ریال

شمارگان: ۱۰۰ جلد

این دوفصلنامه با مجوز شمارهٔ ثبت ۹۱/۳۴۷۱۵ وزارت فرهنگ و ارشاد اسلامی چاپ و منتشر می‌شود.

شاپای‌الکترونیکی: ۰۲۰۴-۲۸۲۱
شاپای چاپی: ۸۹۸۶-۲۸۲۰