

TEFL Journal

TEFL Journal

Three time in a year



Online ISSN:2958- 7549

Print ISSN : 3135-6826



TEFL Journal (TJ)

AL-Afaaf Institute (Ministry of Education)

Vo 4, No 2, 2026





TJ

Online ISSN : 2958- 7549

Print ISSN : 3135-6826

Mobile : 009647734227426

<https://www.tefljournal.com/index.php/TJ>

alafaafinstitute@gmail.com



Editorial Team

Editor-in-Chief

Prof. Z. Ghabanchi

English language and Literature , Faculty of Letters and Humanities
Ferdowsi University of Mashhad , Mashhad , Iran

Editorial secretary

Prof. Ridha Kamil Almusawi, Al-Bayan University , Iraq

Managing Editor

Dr. Taif Alasamy , Ministry of Education, Iraq

Editorial panel

Dr. Kian Pishkar, Department of English language and literature , Islamic Azad University
Jieroft Branch , Kerman , Iran

Prof. Dr. Zeena Abid Ali Dawood Albayati, Mustansiriyah university/college of
Education/department of English, Iraq

Asst. prof. Dr. Raghad khalaf ulaywi, Mustansiriyah university/college of
Education/department of English, Iraq

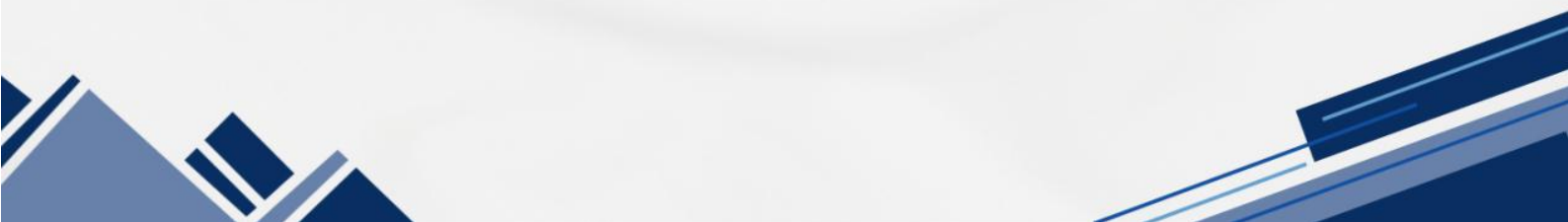
Dr. S. Poornima, Department of English,
Trinity College for Women (Arts & Science), Tamilnadu, India

Dr. K. S. Dhivya, Head, Department of English,
Trinity College for Women (Arts & Science) , Tamilnadu, India

Dr. T. Sridevi, School of Humanities, SRM Institute of Arts and Science, Chennai, India

Prof. Dr. Haydar Jabr Koban, Al-Bayan University , Iraq

Prof. Dr. Chassib Fanukh Abbas, Open Education College in Diwanyah, Iraq



Dr. Armand James A. Vallejo , Department of teacher Education ,
University of Mindanao , Tagum college , Philippine

Prof. Dr. Fatima Rahim, College of Basic Education – Maysan University

Sabriya Said Juma Al Zadejaly , Asst.Lec. Applied Linguistics and TESOL,
Newcastle Upon Tyne University, England

Dr. Ali Jabal Alwan , Islamic University in Diwanyah

Dr. Duygu Şeyma Avşar, Mersin University ,Turkey.

Dr. K. Kaviarasu, Vivekananda College, Agasteeswaram, Kanniyakumari, India

Dr. Riyadh Abbas Ubeid , Al-Salam University, Iraq

Dr. Ahmed Mohammed Bedu, University of Maiduguri, Nigeria

Publication Ethical

Privacy Statement

Privacy Statement:

The names and email addresses submitted on the (TJ) site will be utilized solely for the specified objectives of this journal and will not be disclosed for any other reason or to any third party.

Open Access Statement:

TEFL Journal (TJ) is an open access journal, all of its content is freely accessible to users and their institutions at no cost. Without first requesting permission from the publisher or the author, users are free to read, download, copy, distribute, print, search, or link to the entire texts of the articles, or use them for any other legitimate purpose. This is consistent with the definition of open access provided by the Budapest Open Access Initiative (BOAI). Click here to learn more about BOAI. (<https://www.budapestopenaccessinitiative.org/>)

Plagiarism Policy

The TEFL Journal (TJ) takes the required steps to assess incoming papers for originality, accuracy of the information they include, and proper citation usage.

Publication Ethics and Publication Malpractice Statement:

The TEFL Journal (TJ) is dedicated to upholding the highest moral standards. TJ adheres to publication ethics norms to guarantee that our readers receive a high-caliber academic journal.

Publication Ethics and Publication Malpractice Statement

In recent years, integrity in research publications has become a hotly debated topic. The rise in plagiarism and other types of fraud and misconduct in academic writing is a major worry for the editors of the TEFL Journal (TJ). We see it as a crucial component of our ethical

duties as editors to recognize these difficulties, address them, and devise coping mechanisms. We anticipate that our journal's peer reviewers will be equally conscious. Every research manuscript undergoes editorial internal review and double-blind peer review by a minimum of two reviewers. Similarly, TJ's editorial policy demonstrates its dedication to moral principles. The TJ editors make every effort to guarantee impartial, transparent, and equitable peer review procedures and editorial choices. We will actively pursue any instances of wrongdoing that are found, whether they involve authors, reviewers, or editors. The TEFL Journal (TJ) is dedicated to upholding the highest moral standards. The next section is a comprehensive policy on TJ publication ethics (for authors, editors, and reviewers) and a malpractice statement based on the ethical standards of the Committee on Publication Ethics (COPE).

Duties of Authors:

a) Reporting Principles

The authors of the original work should provide a detailed description of the work done as well as an unbiased analysis on the significance of the study. To enable other researchers to duplicate the study, the researcher has a crucial duty to methodically report on all information pertaining to the data and participants in the study (including the population, sample, and sampling methodology). It is unacceptable to make false or misleading claims that indicate unethical activity.

b) Data Admission and Maintenance

Before or after publication, the raw data for the paper should be ready to be sent to editorial review upon reasonable request. Data should be deposited in a suitable storage area for sharing and further use by others, if permitted by the employer or authors.

c) Novelty, Plagiarism and Concurrent Publication

In order to safeguard the rights of the authors and the journal's reputation against malpractice, TJ takes copyright violations and plagiarism issues extremely seriously and looks for instances of plagiarism or abuse of published papers. Therefore, authors should ensure that their pieces are wholly unique, and if they have used someone else's words or work, they should properly cite or quote that person. Additionally, since TJ does not permit submitting a paper that is being considered or has been published in whole or in part by

another journal, authors should verify that the submitted work is not being considered or accepted for publication elsewhere. The TJ has the authority to retract an article, discuss the issue with the head of the department or the dean of the author's institution, or take appropriate legal action if it is discovered that a paper contains plagiarism, uses third-party material without permission or proper acknowledgment, and the authorship of the work is questioned.

d) Acknowledgment of Sources

It is always necessary to properly credit and acknowledge the works of others. Furthermore, confidential services data (such as reviewing manuscripts or grant applications) and private collected data (such as conversations, letters, and discussions with third parties) must be disclosed with accurate written consent from the source.

e) Paper's Authorship

Only individuals who have contributed significantly to the article—whether by conception, design, execution, or interpretation—may be considered authors. An acknowledgments section should include a list of all other individuals who have contributed significantly to the research effort but do not fit the authorship requirements. The corresponding author and all co-authors must approve the final draft of the work before it is submitted for publication.

f) Hazards, Informed Consent

Manuscripts involving substances, processes, or instruments that could pose any exceptional risks or dangers must specify and identify such hazards. Every document that reports on human research must declare that all procedures were carried out in accordance with applicable laws and institutional rules, and that approval was obtained from the relevant Ethics or suitable Institutional Committee or Committees. When conducting experiments on human beings, the authors must specify in the manuscript whether or not participants gave written or verbal informed consent.

g) Declaration of Conflicting Interest

The Editor must be informed as soon as possible of any financial, commercial, or other fundamental conflicts of interest (such as employment, consulting fees, payments, grants, paid expert testimony, applications/registrations, or other funding) that could influence the article's outcomes or interpretation.

h) Published Works' Significant Errors

In order to amend the article, significant flaws or inaccuracies in published work should be reported as soon as the author becomes aware of them. If the editor or a third party discovers a serious flaw in a published work, it is also the author's duty to quickly retract or amend the work. Additionally, if an erratum, addendum, corrigendum notice, or retraction of the article is judged essential, the authors should collaborate with the editor and publisher.

Duties of Reviewers:

a) Contribution to Editorial Decisions

The core of the scientific method and an important component of scholarly communication is peer review, which helps the editor and authors make timely editorial judgments and improve the manuscript through editorial communications, respectively.

b) Punctuality

Referees or reviewers should inform the editor and withdraw from the review process if they are unable to evaluate a manuscript promptly or do not think they are qualified to do so.

c) Confidentiality

Any manuscripts that are received for review must be handled as confidential papers and cannot be shown, kept, copied, or discussed with others without the head editor's accurate authorization.

d) Objectivity Standards

Reviewers should conduct the review fairly and objectively, and they should clearly state their conclusions without imposing their own.

e) Sources Acknowledgement

The reviewers should identify the proper and appropriate reference of any published work (observation, derivation, argument, discussion, etc.). Additionally, it is the reviewers' duty to notify the editor of any similarities between the article being reviewed and any other known published materials.

f) Disclosure and conflict of interest

The reviewers cannot use any portion of the received materials (published or unpublished) or other information obtained through peer review in their own research or for personal gain without the author's consent to publish the pertinent contribution because all manuscripts received for review must be treated as confidential documents. Additionally, reviewers should refrain from reviewing submissions in which they have financial, competitive, or collaborative conflicts of interest because of their relationships with the authors or institutions.

Duties of Editors:

a) Publication Decisions

The editor of a peer-reviewed journal must promptly determine which of the submitted articles should be considered for publication, working with the editorial board and other relevant societies. The choice is typically based on the work's validation, its significance to readers and scholars, and ethical factors including plagiarism and copyright infringement.

b) Fair Play

The intellectual substance of manuscripts should be the basis for a fair, impartial, and balanced evaluation; the author's origin, nationality, race, gender, sexual orientation, religion, commercial influence, or political ideology should not be taken into consideration.

c) Confidentiality

The corresponding author, reviewers, potential reviewers, editorial advisers, and the publisher are the only parties to whom information on submitted articles may be disclosed, retained, or copied.

d) Disclosure and Conflicts of Interest

The Editors are prohibited from using any portion of accepted materials (published or unpublished) or other information obtained through peer review in their own study or for personal gain because all papers received for evaluation must be treated as confidential documents. The editors should invite a co-editor or another member of the editorial board to study and consider any article in which they have conflicts of interest because of affiliations with the authors or institutions, or because of competitive or collaborative issues. Editors

must also request that all contributors disclose any pertinent interests, and if any conflicting interests are found after publication, they must publicize any changes, retractions, or expressions of concern. Editors should make sure that sponsored submissions go through the same peer-review procedure as normal manuscripts and in accordance with journal guidelines. Only academic merit and reader interest should be taken into account when accepting sponsored papers; commercial motives should not be taken into account.

e) Complaints

When it comes to ethical or conflict-related complaints, the Journal adheres to rational, equitable, and impartial procedures based on the Journal's and Society's policies, when applicable. The editor gets in touch with the author who has been the subject of a complaint. Any objections will have a fair chance to be addressed by the authors. In addition to drawing conclusions from the opposing viewpoint, the editor may make a decision based on the author's response (e.g., if the author is responsive, articulates a significant stance).

f) Involvement and Collaboration in Investigations

The editor should evaluate the ethical issues of a submitted article or published work using suitable and responsive means, such as corresponding with the paper's author or contacting relevant research bodies or institutions. Even if the unethical publishing act is discovered years after publication, every unethical publishing conduct should be taken into consideration and thoroughly investigated (by correction, retraction, expression of concern, etc.).

Duties of Publisher:

The TJ makes sure that the aforementioned requirements for good practice are upheld.

COPE's Guidelines & Flowcharts

TJ is dedicated to adhering to and using the Committee on Publication Ethics' (COPE) rules and flowcharts when reviewing and publishing issues. For additional details, see:

<https://publicationethics.org/>

Procedures for dealing with unethical behavior

Identification of unethical behavior

Anybody should report malpractice and unethical behavior to the editor/publisher at any time. For an investigation to be started, those who alert the editor or publisher about such unethical activity must present sufficient proof.

Investigation

It is the editor's responsibility to start the inquiry and make a decision (consulting or asking the publisher for help, if necessary) without disseminating any accusations to anyone who shouldn't know.

Breaches and Outcomes (in increasing order of severity; may be applied separately or in conjunction)

By educating the author or reviewer about misinterpretation or improper application of permissible norms, minor wrongdoing will be addressed without broader consultation. A formal retraction or withdrawal of a publication from the journal may be granted to the authors.

For a predetermined amount of time, the authors may not be allowed to contribute.

If there is serious misconduct, the accused's employers may be informed or a formal notice outlining the misconduct may be published.

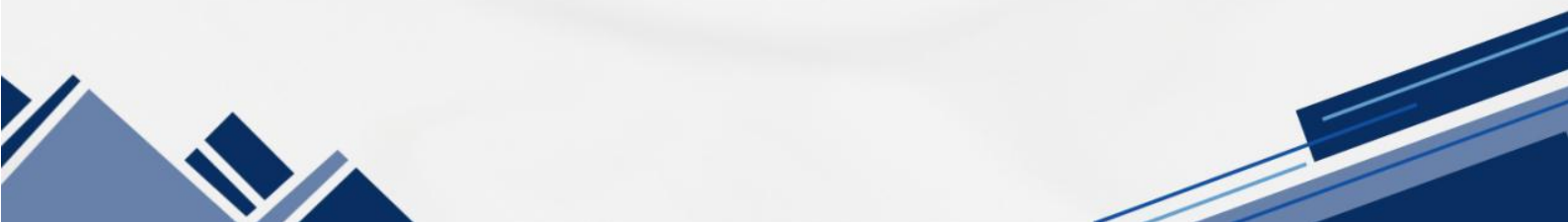
The head of the author's or reviewer's department or funding organization may receive a formal letter.

Reporting the case and outcome to a professional organization or higher authority for further investigation and action.

In any case, the author ought to have a chance to address any accusations. TJ regulations concerning:

Article withdrawal

Articles that represent violations of professional ethical codes, such as multiple submissions, false claims of authorship, plagiarism, falsified data, or are found to be unintentional duplicates of other published articles, are considered in press. These articles have been accepted for publication but have not yet been formally published and will not yet have the full volume, issue, or page information.



Allegations about authorship of contributions

Giving each author credit for their work as a "author" is crucial. Each author must have contributed significantly to the conception, design, execution, or interpretation of the work in order to be considered an author. This is because each author must be accountable for particular aspects of the research. It should be possible for each of the research's authors to openly defend their work. All co-authors must be identified by the author submitting an article.

Concurrent publication or simultaneous submission of duplicates Articles that are submitted for publication must be unique and have never been submitted to another journal. It is not acceptable to submit multiple works to the TJ at the same time. TJ recognizes that artificial intelligence (AI) has the potential to produce valuable content for submissions submitted to the journal for possible publication. The publication does acknowledge the necessity of ensuring the ethical and responsible usage of AI-generated content, though. The journal's rules for using AI-generated content in published articles are outlined in this policy.

Definition of AI-generated content

Any content produced or significantly altered by an AI system is considered AI-generated content for the purposes of this policy. This covers both content that an AI system creates from scratch and content that was first written by a human author but was later significantly altered by an AI system.

When authors employ generative artificial intelligence (AI) and AI-assisted technologies throughout the writing process, they should only do so to enhance language quality and readability. Because AI can produce output that seems authoritative but may be inaccurate, incomplete, or prejudiced, the application of such technology should be done under human supervision and control, and authors should carefully review and revise the outcome.

Principles

After reviewing the AI-generated content, TJ will determine whether it is appropriate for publication if:

- The article accurately describes the AI system utilized to generate the content.

- The content produced by AI must be unique and free of plagiarism.
- The stuff produced by AI should be trustworthy and accurate.
- The content produced by AI does not deceive or mislead readers.

The following details must be submitted to TJ by authors who want to submit manuscripts with AI-generated content:

- A thorough explanation of the AI algorithm that produced the content.
- A duplicate of the initial input data that was used to create the content.
- A copy of the content produced by AI.

The use of generative AI and AI-assisted technologies in the writing process must be disclosed by authors by including a statement at the conclusion of their work in the core manuscript file, prior to the References list. A new section named "Declaration of Generative AI and AI-assisted technologies in the writing process" should contain the declaration. Please be aware that the substance of the work is ultimately the responsibility and accountability of the authors.

Statement: The author or authors used [NAME TOOL / SERVICE] to [REASON] while preparing this work. Following the use of this tool or service, the author or authors took full responsibility for the publication's content and reviewed and edited it as necessary.

Enforcement

Any article that violates this guideline may be rejected or retracted by TJ. Authors who disobey this policy may also face disciplinary action from the journal. The paper will be rejected at any point during the review or publication process, including after publication, if such infractions are found.

Plagiarism Policy

Plagiarism and other types of fraud and misconduct in academic publishing are major concerns for the editorial board of TEFL Journal (TJ). For this kind of plagiarism screening, TJ uses Grammarly software. When plagiarism is found in an article that is submitted for publication in TJ, the following policy is established because TJ recognizes that plagiarism is intolerable:

Definition:

According to the Cambridge Dictionary (2020), plagiarism is the "process or practice of using another person's ideas or work and pretending that it is one's own original work."

Policy:

Papers must be unique, unpublished, and not being considered for publication anywhere else. The TJ has the authority to retract a paper and take appropriate legal action if it is discovered to include plagiarism or to use third-party content without permission or proper acknowledgment, and if the work's authorship is in doubt. When plagiarism is found, the TJ Editor-in-Chief is in charge of reviewing the work and deciding on actions based on the degree of plagiarism found in the work in accordance with the following levels of plagiarism:

Level1:MinorPlagiarism

Without any substantial information or concepts borrowed from the other papers or publications, a brief line or paragraph from another article is considered plagiarism.

Punishment: The authors receive a warning and are asked to make changes to the article and correctly credit the original sources.

Level2:IntermediatePlagiarism

A substantial piece of information, a paragraph, or a sentence in a paper is plagiarized if the original source is not properly cited.

Penalty: The manuscript that was submitted is immediately rejected.

Level3:SeverePlagiarism

A significant amount of a study is plagiarized, which includes a variety of elements such copying ideas, methodologies, and original results (data, formulation, declaration, etc.) from other publications.

Penalty: The authors are prohibited from submitting any more publications to the Journal after the work is automatically rejected.

TEFL Journal (TJ) is dedicated to upholding the highest moral standards. TJ is governed by the publication ethics rules to guarantee that our readers receive a high-caliber academic journal.

Copyright & Licensing

The copyright of papers published in TEFL Journal (TJ) is retained by the authors, who are also free to replicate and distribute their work. The authors give TJ permission to publish the work and claim to be the original publisher. Authors also grant any third party the right to use the article freely as long as its integrity is maintained and its original authors, citation

details and publisher are identified.

TEFL Journal (TJ) is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](http://creativecommons.org/licenses/by-nc/4.0/) (<http://creativecommons.org/licenses/by-nc/4.0/>)(CC BY-NC 4.0). To view a copy of this license, visit: <https://creativecommons.org/licenses/by-nc/4.0/>

Guide for Authors

When an article is submitted to the TEFL Journal (TJ), it is assumed that it has never been published before (apart from as an abstract or as a component of an academic thesis) or that it is not being reviewed for publication elsewhere. This journal will not consider publishing anything that has already been published, copyrighted, or approved for publication. When a manuscript is submitted, it is considered a confirmation that no portion of it is being reviewed by another official publication or is copyrighted by another publisher.

The (TJ) and its Editorial Board are not accountable for the viewpoints stated; instead, it is the duty of the author or authors and their organizations. Before submitting the paper, the relevant author must secure the required copyright release permissions for the use of any copyrighted elements. The authors will receive proofs for approval prior to publication, but the editors retain the right to edit or otherwise modify any contributions. The work's subsequent uses are not the journal's responsibility. If the author wishes to file an infringement lawsuit, it is their obligation.

Authors should make their submissions via the online submission system , so that their papers can be sent for review. Submissions by email are not acceptable and will be discarded.

Please read the instructions below before submitting your story. Please be aware that manuscripts that don't follow these rules might be sent back. Only high-quality manuscripts that fit the objectives and parameters of (TJ) will be examined.

Manuscripts submitted to (TJ) should not have been published before or be considered for publication in another journal.

Format:

The APA style (7th edition), which has the following parts in the following order, should be used for all articles.

Title page: Contains the article's subject, the writers' names, the department or company they work for in English, their address, phone number, and email address.

The title, abstract (150–250 words), keywords (3–5 words), introduction, overview of relevant literature, research methods, findings, and discussion comprise the body. The length of the document should range from 3000 to 8000 words. Acknowledgments: (If required)

References

NOTE: For avoiding desk rejection, all authors seeking for article submission are requested to submit the plagiarism report (Maximum 20%) along with the other required files mentioned above (i.e., title page & main manuscript, ...) to the Journal. They must get the legal plagiarism report from the links below:

www.ithenticate.com

www.turnitin.com

www.unicheck.com

Article Processing Charge : Free

Submission Preparation Checklist:

Authors must confirm that their proposal complies with each of the following requirements as part of the submission process. Submissions that don't follow these rules won't be

Table of contents

Professional Distance Theory (PDT) in Iranian EFL High School Classrooms: A Framework for Managing Student Exploitation of Teacher Goodwill.....	14
---	-----------

Athraa Abd Ali Lateef

Introducing Effort Atrophy Theory through a Comparative Study of Iranian and Iraqi EFL High School Students.....	34
---	-----------

Zahraa Mohammed Hussein

The Effect of Multimodal Instruction on Vocabulary Learning and Retention in Iraqi EFL Classrooms.....	54
---	-----------

Rusul Esmaeel Sakran



Professional Distance Theory (PDT) in Iranian EFL High School Classrooms: A Framework for Managing Student Exploitation of Teacher Goodwill

Athraa Abd Ali Lateef

Article Info

Article History

Received:

1 March 2026

Revised

1 May 2026

Accepted:

1 July 2026

Published

25 July 2026

Keywords

Professional Distance Theory (PDT), teacher-student relationship, high-power-distance culture, classroom management, student exploitation

Abstract

This study develops and empirically tests Professional Distance Theory (PDT), which proposes that in Iran's high-power-distance and collectivist culture, an optimal teacher-student distance, neither overly close nor excessively authoritarian, is necessary to prevent students from exploiting teacher goodwill and to maintain effective classroom management. Data collection involved 150 Iranian EFL high school teachers surveyed using the Teacher Professional Distance Scale (TPDS) and semi-structured interviews within a sequential explanatory mixed-method design. Quantitative findings revealed that teachers reporting closer proximity also reported significantly greater instances of classroom norm violations, including grade negotiation, favoritism-seeking, and manipulation of teacher flexibility. In contrast, teachers maintaining moderate professional distance reported lower exploitation rates and higher perceived authority without sacrificing student engagement. Qualitative evidence underscored three primary mechanisms: (1) the Kindness Trap (students learn to interpret closeness as weakness to be exploited), (2) Role Ambiguity (blurred boundaries give rise to increased testing behaviors), and (3) Authority Erosion (excessive friendliness undermines perceived legitimacy). The findings support a four-level Professional Distance Continuum, Distant, Moderate Professional Distance, Personal, and Overinvolved, with Moderate Professional Distance identified as most effective. These results challenge the universal application of Western relational models and provide a culturally situated framework for teacher professional development in non-Western hierarchical education systems.

1. Introduction

The teacher-student relationship is widely recognized as an important factor influencing classroom climate, student involvement, and academic performance (Hamre & Pianta, 2006; Eccles & Wigfield, 2002). From a Western educational psychology perspective, particularly within Self-Determination Theory (SDT), close, warm, and supportive teacher-student relationships have been consistently linked to greater intrinsic motivation, autonomy, and well-being (Ryan & Deci, 2017; Reeve, 2009). The dominant school-based paradigm in the educational literature advocates for reduced relational distance (Niemic & Ryan, 2009), less teacher-student aloofness, and more emotionally supportive and collaborative classroom cultures.

However, the universality of this framework has been increasingly challenged. Lee and Lee (2018) found that teachers' cultural dispositions toward uncertainty and hierarchical relationships in contexts like Singapore significantly moderate the effectiveness of collaborative practices, indicating that approaches grounded in low-power-distance values may not translate effectively to high-power-distance societies. According to Hofstede's (2001) cultural dimensions model, Iran is characterized as a high-power-distance culture, with hierarchical relationships, deference to authority, and formal social roles being strongly rooted in institutional practice. Observational studies confirm these patterns: in Iranian classrooms, professors tightly control discourse and student contributions, revealing high power distance and uncertainty avoidance (Moradi, 2024).

A comparative study of EFL and non-EFL teachers in Iranian high schools revealed that English teachers demonstrated statistically significantly higher power distance toward their students than non-English teachers, suggesting that the language learning environment may exacerbate hierarchical dynamics (Bijani et al., 2022). Despite this cultural complexity in EFL classrooms, the existing literature provides relatively few accounts of effective practice in high-power-distance EFL contexts. Iranian EFL instructors operate in a particularly challenging professional context characterized by competitive institutional environments, exam-driven curricula, and cultural norms that reinforce teacher authority in ways that differ from Western settings (Aliakbari, 2015; Abdoli et al., 2017).

Most importantly, many Iranian teachers experience a troubling phenomenon: when they attempt to adopt Western models of warm, friendly relationship-building, this goodwill is often exploited by students. Students interpret teacher flexibility, leniency, or relational kindness as opportunities to negotiate grades, request special favors, circumvent rules, or undermine classroom authority. According

to Abdoli et al. (2017), Isfahan students perceive teacher favoritism as the most frequent barrier to building positive relationships, while teachers identify inconsistencies between their stated principles and actual teaching behavior as a source of student dissatisfaction. This process of misuse of teacher goodwill in high-power-distance, collectivist cultures has received little scholarly attention.

The primary variable in this research is professional distance, defined as the level of formality, role clarity, and relational boundary maintenance employed by teachers in relation to students. Secondary variables include student exploitation acts (e.g., grade negotiation, rule challenging, inappropriate requests, boundary testing) and classroom authority (defined as the perceived capacity of teachers to maintain classroom rules, discipline, and instructional order). The overall goal of this work is to provide an empirical foundation for PDT, which argues that in collectivist cultures with high power distance, such as Iran, establishing an appropriate level of relational distance is necessary to prevent exploitation of teacher goodwill and to promote effective classroom management.

2. Literature Review

2.1 Theoretical Background

2.1.1 SDT and Universality

According to SDT (Deci & Ryan, 1985), fulfilling three fundamental psychological needs, autonomy, competence, and relatedness, cultivates intrinsic motivation and well-being across cultures (Ryan & Deci, 2017). Relatedness, the need to feel connected and cared for, has been articulated in educational research as teacher closeness and emotional support for students (Niemic & Ryan, 2009). However, the relational prescriptions articulated by SDT may not be universally applicable across cultural contexts. In Asian and Middle Eastern contexts, some researchers suggest that teacher control may be even more important than autonomy support, and that relatedness may be expressed differently across cultures (Reeve et al., 2014). The need for autonomy may be universal, but its expression through support is culturally constructed; teachers from high-power-distance cultures may view autonomy support as inappropriate to their role or as undermining their legitimate authority. Therefore, uncritically employing Western relational models may not lead to relatedness satisfaction but rather to role confusion and degradation of authority.

2.1.2 Hofstede's Power Distance and Collectivism in the Iranian Context

Hofstede (2001) categorized Iran as a high-power-distance culture, characterized by acceptance of hierarchical order, respect for authority, and expectations of formal role differentiation. While

collectivism, another Iranian cultural dimension, emphasizes group harmony, mutual loyalty, and relational bonds among group members, these values do not necessarily translate to egalitarian teacher-student dynamics. In fact, Moradi (2024) showed that Iranian teachers enacted more power inequity than their Chinese counterparts, with professors micromanaging classroom discourse. Critically, power distance moderates student expectations of teacher behavior: high-power-distance learners expect clear markers of authority, formal role enactment, and dominance. When teachers violate these norms through overly friendly or egalitarian behavior, students perceive this not as support but as role failure, creating potential for boundary testing and exploitation.

2.1.3 Teacher Authority and Boundary Theory

Teacher authority refers to the teacher’s legitimate right to direct student behavior, enforce rules, and assess student performance. Authority can derive from positional status, expertise, or relational quality. In high-power-distance cultures, positional authority is paramount, and students expect teachers to behave in ways that reinforce this authority. Practices that blur boundaries, such as sharing excessively personal information, deferring to inappropriate requests, or failing to enforce consistent consequences, undermine positional authority while inviting exploitation. Boundary management, therefore, is not merely a professional ethics issue but a functional necessity for effective classroom governance in hierarchical systems.

2.1.4 Hypothesis on Kindness Exploitation

Expanding upon attribution theory and social exchange theory, the kindness exploitation hypothesis suggests that in high-power-distance cultures, relational closeness may not simply be interpreted as helpfulness by students but may be cognitively reframed as a resource to be manipulated. When teachers are “too nice,” students may attribute this behavior to teacher weakness, a desire for validation, or lack of alternatives, perceptions that can directly encourage manipulative behavior. Paradoxically, the relational strategies recommended by Western educational psychology may become self-defeating in high-power-distance contexts.

Table 1: Professional Distance Continuum			
Level	Label	Characteristic	Predicted Outcome
1	Distant/Authoritarian	Strict rules, minimal interaction, frequent punishment, no personal sharing	Low exploitation; low engagement; student resentment

2	Moderate Professional Distance	Clear rules, consistent enforcement, appropriate personal sharing, role clarity, respectful communication	Low exploitation; moderate-to-high engagement; stable authority
3	Personal/Overly Close	Frequent personal sharing, high flexibility, minimal consequences, friendship expectations	High exploitation; unstable authority; student manipulation
4	Overinvolved/Inappropriate	Blatant favoritism, secrecy, emotional dependence, boundary violation	Extreme exploitation; authority collapse; ethical violations

2.2 Prior Related Studies

Iranian EFL classroom management literature has reported on assertive strategies centered on organization, teacher-student relationship management, and consequences in the Iranian context (Aliakbari, 2015). Iranian teachers employ various classroom management strategies, including organizational teaching methods, teacher-student relationship approaches, and reward-punishment systems, with teacher assertiveness positively correlated with student performance. Importantly, assertiveness in the Iranian context may not involve relational closeness but rather consistent and clear communication of expectations and enforcement of consequences, elements of professional distance.

Regarding student exploitation, the main barriers to constructing positive relationships identified from the student perspective were teacher favoritism and inappropriate teacher responses to students with achievement and discipline issues (Abdoli et al., 2017). Teachers themselves cited obstacles such as classroom boredom, contradictions between their speech and action, and making students hopeless about pursuing their education. These perceptions of favoritism (a predictable result of differential closeness) compromise teacher credibility and institutional legitimacy, while inconsistency between words (showing caring) and actions (failing to enforce consequences) creates exactly the role ambiguity that allows behavioral manipulation and testing.

Discursive enactment of power in Iranian EFL classrooms has been analyzed through critical discourse analysis, showing that teacher-dominated instructional environments can be characterized as

oppressive educational practices from a critical perspective. However, what is characterized as oppression can also be considered an enactment of culturally normative authoritative instruction.

Critically, Western-style egalitarianism may inadvertently create a power vacuum that students fill through testing, negotiation, and manipulation, dynamics that teachers subjectively experience as exploitation of their goodwill. Despite extensive research on classroom management and teacher-student relations in Iran, no investigation has systematically theorized the link between professional distance and student exploitation behavior. Iranian teachers find themselves caught between Western models of teacher-student intimacy and the real socio-cultural conditions of high-power-distance contexts with conflicting signals about their role.

This gap is addressed by the current study, which proposes Professional Distance Theory (PDT) and empirically tests its core assumptions. Although considerable research on classroom management and teacher-student relationships in Iran has been undertaken, three major areas require attention: (1) the need for a systematic theory establishing student exploitation of teacher goodwill as a function of professional distance in high-power-distance cultures; (2) a lack of empirical studies measuring how teacher relational proximity affects the frequency of boundary violations with students in Iranian EFL contexts; and (3) a dearth of culturally situated models for managing professional distance that balance maintaining authority with promoting student engagement. This study seeks to fill these gaps by investigating the following research questions:

- **RQ1:** What are the reported forms of student exploitation of teacher goodwill among Iranian EFL high school teachers?
- **RQ2:** Is teacher professional distance (relational proximity) related to the frequency of student boundary violations?
- **RQ3:** What is the ideal extent of professional distance perceived by teachers to prevent exploitation while still stimulating student engagement?
- **RQ4:** What are the main elements and propositions of Professional Distance Theory (PDT) based on the data?

The importance of this research is multi-faceted. Theoretically, it presents PDT as a culturally situated alternative to universalist Western relational models by articulating a direct connection between professional distance and student exploitation behavior. Empirically, it offers practical guidelines for Iranian EFL teachers on establishing optimal professional distance that allows them to maintain high-quality classroom climate and student motivation. Drawing out the mechanisms through which goodwill

is exploited may enable this study to inform teacher professional development and classroom management training not only in Iran but also in similar Middle Eastern and Asian contexts.

3. Methodology

3.1 Research Design

A sequential explanatory mixed-method design was employed. The quantitative phase used a cross-sectional survey to measure professional distance, student exploitation frequency, and classroom authority across a large sample. In the qualitative phase, semi-structured interviews were conducted with a subset of participants to investigate mechanisms, contextual factors, and teacher strategies. This design was selected to allow for both broad generalizable claims (quantitative) and extensive contextual insights (qualitative) into the complex culturally embedded phenomenon.

3.2 Participants

The sample consisted of 150 Iranian EFL high school teachers (86 female, 64 male) from 25 public and private schools in Tehran, Isfahan, and Mashhad provinces, selected using stratified random sampling. Stratification was conducted by school type (51% public, 49% private), sex, and teaching experience (mean = 9.2 years, SD = 5.1). Inclusion criteria required that participants: (1) currently teach English at the high school level, (2) have taught in their current school for at least three years, and (3) work full-time. Private school teachers were purposefully oversampled to reflect differences in school-level autonomy and systemic pressure. Informed consent was obtained from all participants.

3.3 Instrument

The researcher-developed Teacher Professional Distance Scale (TPDS) was designed for this study with 28 items across four subscales: (1) Professional Distance (10 items; e.g., “I maintain clear emotional boundaries with my students”), (2) Student Exploitation Frequency (8 items; e.g., “Students frequently try to negotiate for higher grades”), (3) Rule Enforcement Consistency (5 items; e.g., “I enforce consequences consistently regardless of student popularity”), and (4) Role Clarity (5 items; e.g., “Students understand that I am their teacher, not their friend”). Responses were collected on a 5-point Likert scale (1 = never/strongly disagree; 5 = always/strongly agree).

The TPDS was designed in English, translated to Persian using a blind back-translation procedure, and piloted with 30 teachers, showing a Cronbach’s alpha of 0.91. The semi-structured interview protocol included 12 open-ended questions designed to gauge teacher experiences with student

exploitation, coping strategies for boundary management, perceived cultural expectations, and perceived effectiveness of different distance levels. All interviews were conducted in Persian, audio-recorded, transcribed, and translated into English for subsequent analysis.

3.4 Data Collection Procedure

Data were collected over a three-month period (January through March 2025). At each participating school, research assistants distributed the TPDS paper questionnaire during teacher meetings; completion took approximately 15 minutes. Participation was voluntary and conducted anonymously. A total of 160 questionnaires were administered; 150 were returned fully completed (93.8% response rate).

Subsequently, 25 teachers (15 from the quantitative sample) were purposively selected for semi-structured interviews based on TPDS scores representing low, moderate, and high professional distance levels. Interviews lasted 30–45 minutes each and were conducted in private rooms after school hours.

3.5 Data Analysis Methods

Quantitative data were analyzed using IBM SPSS v.29. Preliminary procedures included descriptive statistics, reliability testing (Cronbach's alpha), and normality analysis (Shapiro–Wilk). For RQ1, descriptive frequencies and means for exploitation forms were determined. For RQ2, Pearson correlation coefficients were calculated between professional distance scores and student exploitation frequency. ANCOVA was used to examine exploitation frequency by teacher age, experience level, and institution type. For RQ3, mean comparisons identified the professional distance level associated with minimum exploitation and highest authority ratings; one-way ANOVA and Tukey HSD post hoc tests were conducted.

Qualitative data were analyzed using thematic analysis following Braun and Clarke's (2006) six-phase procedure. Transcripts were independently coded by two coders; inter-coder reliability exceeded 90%.

4. Results

4.1 Preliminary Analysis

The Shapiro–Wilk test indicated that all variables were normally distributed ($p > 0.05$). The TPDS demonstrated a Cronbach's alpha of 0.91 (subscale alphas ranged from 0.84 to 0.89). Descriptive statistics for key variables are outlined in Table 2.

Table 2: Descriptive Statistics for Key Variables

Variable	M	SD	Range	Skewness
Professional Distance	3.42	0.71	1.8–4.9	-0.24
Student Exploitation Frequency	3.18	0.92	1.2–5.0	0.17
Rule Enforcement Consistency	3.56	0.65	1.9–5.0	-0.43
Role Clarity	3.79	0.58	2.2–5.0	-0.58

Note: N = 150. All scales range from 1 to 5.

4.2 Results for Research Question 1

The first research question addressed reported forms of student exploitation of teacher goodwill among Iranian EFL high school teachers. Four distinct exploitation types were identified, with frequencies presented in Table 3. Grade negotiation (M = 3.89, SD = 0.94) was the most widely reported form of exploitation, followed by rule challenging (M = 3.45, SD = 0.88), inappropriate requests (M = 3.21, SD = 0.78), and boundary testing (M = 2.87, SD = 0.72).

Table 3: Frequency of Student Exploitation Behaviors

Exploitation Type	M (SD)	% Reporting Often/Always
Grade Negotiation	3.89 (0.94)	67.3%
Rule Challenging	3.45 (0.88)	54.7%
Inappropriate Requests	3.21 (0.78)	42.7%
Boundary Testing	2.87 (0.72)	29.3%

Grade negotiation (M = 3.89) was the most common and severe form of exploitation, with a majority (67.3%) of teachers reporting that students asked them to raise their grades without justification. Rule challenging was also common (54.7%), signifying that students challenged teacher authority when they felt relational closeness (e.g., questioning why they must follow rules that others do not). More than 40% of teachers (42.7%) reported inappropriate requests (e.g., requesting special treatment based on personal relationships). Finally, boundary testing (29.3%) was the least common but still notable, with nearly 30% of teachers reporting that students frequently tested behavioral boundaries.

4.3 Results for Research Question 2

The second research question sought to test whether a relationship exists between teacher professional

distance (relational proximity) and the frequency of student boundary violations. Pearson correlation analysis revealed a statistically significant negative correlation between professional distance and frequency of student exploitation ($r = -0.58$, $p < 0.001$, 95% CI $[-0.68, -0.46]$), suggesting that teachers with higher professional distance experienced significantly lower rates of student exploitation.

Table 4: Correlation Matrix of Key Variables

Variable	Professional Distance	Exploitation Frequency	Rule Consistency	Role Clarity
Professional Distance	1.00			
Exploitation Frequency	-0.58**	1.00		
Rule Consistency	0.51**	-0.44**	1.00	
Role Clarity	0.46**	-0.39**	0.48**	1.00

$p < 0.01$ (two-tailed)

In the ANCOVA controlling for teacher age and school type, no significant effect of these covariates was found ($F(1, 146) = 0.84$; $p = 0.36$). The relationship between professional distance and exploitation frequency was uniform across school types (public vs. private) and teacher experience levels. The strong negative correlation indicates that student exploitation behavior decreased dramatically as teachers moved from closer, more informal relationships to greater professional distance. The magnitude of the association ($r = -0.58$) reflects that professional distance explains approximately 34% of the variance in exploitation frequency, a considerable effect. The lack of covariate influence suggests that this pattern holds true across all teaching contexts in Iran.

4.4 Results for Research Question 3

The third research question sought to define the degree of professional distance perceived by teachers as most effective for avoiding exploitation while maintaining student engagement. ANOVA comparing distance quartiles revealed significant differences in exploitation frequency: $F(3, 146) = 31.47$, $p < 0.001$, $\eta^2 = 0.39$, indicating that distance level explains 39% of the variance in exploitation. Tukey HSD post hoc tests showed that teachers in the moderately distant quartile (Professional Distance Score 3.8–4.2) displayed the lowest exploitation prevalence ($M = 2.34$, $SD = 0.61$) and the highest classroom authority ratings ($M = 4.23$, $SD = 0.54$). Teachers in the closest quartile (very close relationships) showed the highest exploitation rates ($M = 4.21$, $SD = 0.78$) and the lowest authority ($M = 2.56$, $SD = 0.69$).

Table 5: Professional Distance in Use: Exploitation and Authority

Distance Level	n	Professional Distance M (SD)	Exploitation M (SD)	Authority M (SD)
Very Close (Quartile 1)	38	2.41 (0.34)	4.21 (0.78)	2.56 (0.69)
Moderately Close (Quartile 2)	42	3.18 (0.29)	3.45 (0.71)	3.34 (0.58)
Moderately Distant (Quartile 3)	40	3.94 (0.18)	2.34 (0.61)	4.23 (0.54)
Distant/Authoritarian (Quartile 4)	30	4.63 (0.22)	2.89 (0.67)	3.87 (0.62)

The optimal distance level was not the maximum possible distance (authoritarian quartile 4), which was associated with increased exploitation and lower authority compared to the moderately distant level. Instead, the “sweet spot” was moderate professional distance: clear boundaries, consistent expectations, appropriate role differentiation, and polite communication, but without excessive personal sharing or the inflexibility that can lead to student resentment. The curvilinear pattern indicates that both overly close and overly authoritarian approaches yield worse outcomes.

4.5 Results for Research Question 4

The fourth research question aimed to identify the components and propositions of PDT derived from the data. Qualitative thematic analysis of interview transcripts (n = 25) revealed three main mechanisms connecting low professional distance to student exploitation and five core propositions of PDT.

Table 6: Qualitative Mechanisms of Student Exploitation

Mechanism	Description	Frequency (%) of Participants)
Kindness Trap	Students perceive teacher closeness as weakness to be exploited	92%
Role Ambiguity	Blurred boundaries encourage testing of teacher limits	84%
Authority Erosion	Excessive friendliness erodes perceived legitimacy	88%

The Kindness Trap is a counterintuitive phenomenon in which students reinterpret teacher flexibility, leniency, or personal warmth not as genuine support but as a sign of weakness. In high-

power-distance cultures like Iran, students expect clear hierarchical boundaries. When teachers deviate from this expectation by being “too nice,” they implicitly signal that their authority and power are subject to challenge or that rules can be renegotiated. Consequently, students take advantage of this perceived weakness by escalating demands, making special requests, or testing teacher tolerance.

“The more flexible I was, the more they pushed. They saw every accommodation as a sign they could demand more.”

This quotation vividly describes the Kindness Trap. The teacher describes a clear dose-response pattern: each act of flexibility was followed by increased student pressure. The phrase “saw every accommodation as a sign they could demand more” suggests that students actively monitored teacher behavior and viewed concessions as entitlements rather than gifts. Teacher goodwill was systematically exploited because the students’ cultural frame did not recognize unilateral leniency as legitimate teaching practice but rather as a resource to be extracted.

Role Ambiguity occurs when the line between teacher and student begins to fade, often due to informality, self-disclosure, or inconsistent enforcement. In hierarchical educational cultures, students need clear role definitions to determine their behavior and what they may expect from teachers. When these distinctions become blurred, students experience confusion that they often resolve by actively testing boundaries. Ambiguity itself becomes an invitation to probe how far they can bend the rules.

“They started calling me by my first name, then asked me to overlook their absences. The line became fuzzy.”

This quote demonstrates a clear sequence of escalation. Changing address (from formal title to first name) was the first sign of boundary erosion. After the teacher accepted this informality, students logically followed with substantive demands (overlooking absences). The teacher’s closing remark, “The line became fuzzy,” acknowledges that the teacher could no longer distinguish between appropriate and inappropriate requests. This fuzziness is precisely what makes exploitation possible: when rules are ambiguous, students feel entitled to negotiate.

Authority Erosion refers to the decline in a teacher’s perceived authority to direct behavior, enforce rules, and evaluate performance. High-power-distance cultures maintain authority through role-consistent behavior (i.e., clear boundaries, professional formality, and consistent consequences). When teachers adopt overly friendly or egalitarian approaches, students adjust their expectations accordingly. Later, if the same teacher attempts to assert authority (e.g., by imposing rules or consequences), students perceive this as inconsistent or hypocritical. Teacher authority collapses

because it was never properly established.

“When I tried to be strict after being friendly, they laughed. I had lost my authority permanently.”

This quote illustrates the irreversible nature of authority erosion. The teacher’s shift from friendly to strict was met not with compliance but with ridicule (“they laughed”). Laughter here is not merely an emotional reaction; it is a social signal that the teacher is no longer taken seriously as an authority figure. The phrase “lost my authority permanently” suggests that the teacher believes this loss cannot be recovered within that class. The implication is that professional distance must be established from the outset; attempting to introduce it after a period of closeness is ineffective.

Table 7: Core Propositions of Professional Distance Theory (PDT)

Proposition	Description
P1	In high-power-distance cultures, teacher relational closeness is associated with increased student boundary-testing behavior.
P2	The relationship between professional distance and student exploitation is curvilinear; both overly close and overly authoritarian distances produce suboptimal outcomes.
P3	Moderate professional distance (clear boundaries, consistent enforcement, role-appropriate communication) optimally balances exploitation prevention and student engagement.
P4	Professional distance is not merely authoritarianism; closeness without distance predicts stable authority collapse.
P5	Cultural moderators (power distance, collectivism) determine the optimal distance threshold; PDT is culturally situated, not a universal framework.

Qualitative results across all 25 interviews strongly supported the findings. The Kindness Trap was almost universally applicable (92%), with participants frequently reporting how attempts to be nice backfired. Teacher after teacher described losing authority after initial overfriendliness. Role Ambiguity (84%) and Authority Erosion (88%) were also highly prevalent, suggesting systematic rather than anecdotal mechanisms. The five PDT propositions together create a cohesive, empirically grounded theory of professional distance in high-power-distance EFL contexts.

5. Discussion

5.1 Discussion of Research Question 1

The first research question explored the extent to which Iranian EFL high school teachers reported student exploitation of teacher goodwill and the forms this exploitation took. The results indicate that Iranian EFL teachers frequently reported student exploitation behaviors and were most likely to encounter grade negotiation (67.3%), rule challenging (54.7%), and inappropriate requests (42.7%).

The finding that grade negotiation was the most prominent form of exploitation is consistent with qualitative reports from Abdoli et al. (2017), where teacher favoritism was identified as a predominant obstacle to building teacher-student relationships from the student perspective. The high frequency of grade negotiation reveals that Iranian students actively utilize relational proximity to pursue substantive academic benefits, a strategic action made possible when teachers do not maintain clear evaluative boundaries.

Notably, the regularity of rule challenging is striking. From the perspective of attribution theory, when teachers enact relational closeness, students tend to attribute teacher authority to personal whim rather than to legitimate positional authority, leading students to actively question and negotiate rules. Social exchange theory similarly predicts that students who are close with their teachers perceive an implicit exchange system: their loyalty and compliance are traded for preferential treatment. This becomes particularly problematic when there is inconsistent enforcement of rules built into the implicit contract, which can lead to frustration and increased challenging behavior overall.

5.2 Discussion of Research Question 2

The second research question examined whether a meaningful relationship exists between teacher professional distance (relational proximity) and student boundary violations. A significant negative correlation was found ($r = -0.58$) between professional distance and frequency of student exploitation, with professional distance accounting for approximately 34% of the variance. This finding provides direct support for PDT Proposition 1: teacher relational closeness is associated with increased student boundary-testing behavior in high-power-distance cultures.

While the relatedness construct in Self-Determination Theory suggests universal need satisfaction, the Iranian context, characterized by high power distance and role differentiation, generates different expectations. Students demand that teachers uphold formality; when teachers fail to meet this expectation, cognitive dissonance may be alleviated by interpreting closeness as a sign of weakness

rather than strength. Approaching the analysis from an attributional perspective, students in high-power-distance cultures are predisposed to attribute teacher behavior to role-consistent traits such as authority and expertise. By enacting egalitarian closeness, teachers act in ways inconsistent with student expectations, leading to situational attributions that diminish teacher legitimacy.

The correlation estimate ($r = -0.58$) is higher than commonly considered effect sizes for classroom management variables, indicating that professional distance is especially influential in predicting student behavior. According to Aliakbari (2015), Iranian teachers who utilized assertive management strategies (i.e., clear communication and consistent consequence enforcement, elements of professional distance) achieved higher student performance. The current study expands on this finding by specifically comparing distance to the prevention of negative behavior, as opposed to the generation of positive outcomes.

5.3 Discussion of Research Question 3

The third research question sought to determine the degree of professional distance that teachers consider optimally efficient for preventing exploitation while maintaining student engagement. The most effective level was moderate professional distance (Quartile 3), not maximum distance; both close relationships and authoritarian relationships were associated with increased exploitation. This curvilinear relationship is a fundamental feature of PDT and is directly explicated in Proposition 2.

The finding that authoritarian practice (Quartile 4) was associated with more exploitation than moderate distance may seem counterintuitive. However, excessive authoritarianism, harsh punishment, inflexibility, and lack of explanation for rules, can lead to student resistance and covert boundary testing. In authoritarian classrooms, students may not challenge rules outright but may exploit them in more subtle ways: withholding effort, pretending to comply, or working around loopholes. By contrast, moderate distance combines clear role boundaries with respectful communication, consistent consequence enforcement, and appropriate personal sharing. This preserves teacher positional authority while suggesting that rules are sensible and just. Lee and Lee (2018) discovered that teachers' cultural dispositions toward uncertainty management played a significant moderating role in the effectiveness of collaborative practice in hierarchical cultures; the moderate PDT position provides precisely the structure that high-uncertainty-avoidance students require.

5.4 Discussion of Research Question 4

The fourth research question focused on presenting the main elements and propositions of PDT based on

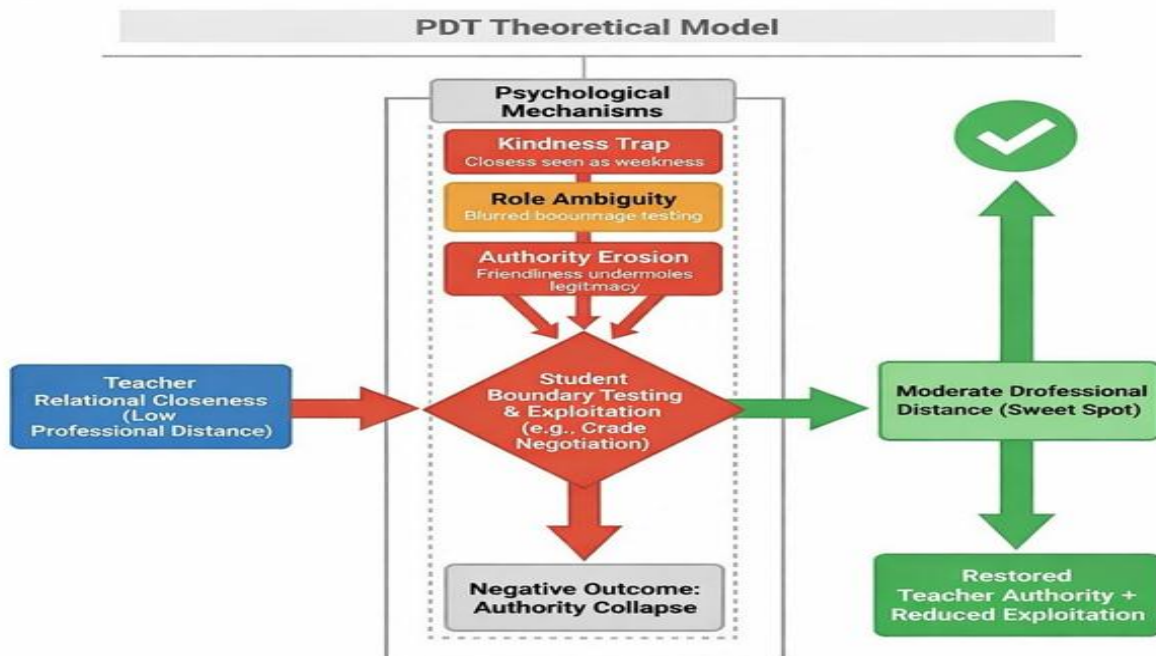
the data. Three fundamental mechanisms (Kindness Trap, Role Ambiguity, and Authority Erosion) were discovered, and five main principles of PDT were conceived.

The Kindness Trap (P1) exemplifies PDT's core contradiction: an act believed to be one of goodwill (flexibility, leniency, informality) is perceived by students as an exploitable resource. This mechanism is consistent with social exchange theory: when teachers show unilateral generosity without requiring reciprocal responsibility from students, students may develop a sense of entitlement rather than gratitude.

Role Ambiguity (P3) is supported by cross-cultural research showing that students in high-power-distance cultures experience anxiety when role boundaries are ambiguous. As a psychological necessity, role clarity is not merely an administrative nicety; without it, testing behavior emerges.

Authority Erosion (P4) aligns with teacher proximity continuum frameworks demonstrating that overly close teacher behavior is significantly related to low professional credibility. The most actionable insight from this theory is the proposition that moderate distance optimizes outcomes (P2).

Figure 1: PDT Model



The Professional Distance Continuum (Table 1) and PDT Model (Figure 1) serve as practical foundations for teacher self-assessment and professional development. The claim that PDT is culturally situated (P5) is an essential qualification: PDT does not assert universal relevance but specifies cultural moderators (power distance, collectivism, uncertainty avoidance) relevant for defining optimal distance thresholds in specific contexts.

6. Conclusion

6.1 Summary of Conclusions

This study developed PDT and tested it empirically, demonstrating that maintaining moderate professional distance, as opposed to the Western model of close egalitarian relationships, in the high-power-distance Iranian EFL context optimizes classroom authority while minimizing student exploitation of teacher goodwill. Key conclusions of the study are:

1. Iranian EFL teachers report high levels of student exploitation behaviors, including grade negotiation and rule challenging.
2. Professional distance is strongly negatively correlated with the frequency of exploitation.
3. Moderate professional distance, rather than extreme positions, produces optimal outcomes.
4. Three mechanisms (Kindness Trap, Role Ambiguity, Authority Erosion) and five PDT propositions create a cohesive cultural theory.

6.2 Research Implications

PDT challenges the universality of Western relational models across different contexts and presents a culturally situated alternative for high-power-distance contexts. It extends the relatedness concept of Self-Determination Theory by arguing that relatedness may be maximized in these contexts not through closeness but through role clarity, respect, and fair consistency. In high-power-distance cultures, professional distance (rather than relatedness in the Western sense) may be the primary means by which students' psychological needs are met; such students require role security and formal authority as well as emotional support.

Teacher training in Iran should explicitly address the Kindness Trap and alert teachers to the fallacy that Western relational models are readily transferable. Pre-service and in-service programs can use the Professional Distance Continuum to help teachers identify their current positioning. Training should emphasize that moderate distance, clear boundaries, consistent enforcement, appropriate personal sharing, is not coldness but professional respect.

School leaders should examine policies that coerce teachers toward excessive closeness (e.g., “be a friend to your students” mandates). In light of professional distance appropriateness, evaluation criteria should be revised to account for the extent to which teachers model care while maintaining appropriate boundaries. PDT offers a framework for teachers in international contexts to align their relational styles with their students' cultural expectations, rather than being dictated by models from other cultures. This

is particularly important for expatriate teachers working in Iran and similar high-power-distance settings.

6.3 Limitations of the Study

Several limitations should be acknowledged:

1. Student perspectives were not collected, as the study relied exclusively on teacher self-report. This is an interpretive constraint that narrows the theoretical claims considerably.
2. The cross-sectional design precludes causal inferences regarding the direction of the professional distance–exploitation relationship. Associations should be interpreted as correlational rather than causal.
3. While stratified, the sample may not fully represent all Iranian provinces or educational levels.
4. Social desirability bias may have affected reporting, particularly on sensitive issues like favoritism.
5. The TPDS, while demonstrating strong internal reliability ($\alpha = 0.91$), requires further validation evidence. No evidence of construct, convergent, or discriminant validity is currently available. The instrument should therefore be treated as preliminary pending cross-validation with independent samples.

6.4 Future Directions for Further Research

Future research should:

1. Examine longitudinal relationships to investigate causal directions.
2. Obtain student perspectives on perceived professional distance and the distance levels that motivate them.
3. Test PDT in other high-power-distance contexts (e.g., Iraq, Saudi Arabia, Turkey, China) to determine generalizability.
4. Compare public versus private schools to evaluate distance dynamics.
5. Develop and evaluate professional distance training interventions.
6. Examine potential interactions among teacher gender, student gender, and professional distance appropriateness.

References

- Abdoli, A., Zeinedini Meimand, F., & Ayati, M. (2017). Obstacles to establishing relationships in the teaching-learning process from the viewpoint of high school teachers and students in Isfahan. *Quarterly Journal of Family and Research*, 14(2), 7–30.
- Aliakbari, M. (2015). Assertive classroom management strategies and students' performance: The case of EFL classrooms. *Cogent Education*, 2(1), 1–12.
- Bijani, H., Jamali, S., & Orabah, S. S. B. (2022). A comparative study of power distance of English teachers and non-English teachers in classroom interaction in Iranian high schools. *Academic Journal of Interdisciplinary Studies*, 11(2), 206–218.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum Press.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53, 109–132.
- Hamre, B. K., & Pianta, R. C. (2006). *Student-teacher relationships*. National Association of School Psychologists.
- Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations* (2nd ed.). Sage Publications.
- Lee, D. H. L., & Lee, W. O. (2018). Transformational change in instruction with professional learning communities? The influence of teacher cultural disposition in high power distance contexts. *Journal of Educational Change*, 19(4), 463–488.
- Moradi, E. (2024). Examining Hofstede's cultural dimensions in Iranian and Chinese contexts: A mixed methods approach. *Innovare Journal of Education*, 12(2), 6–14.
- Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and Research in Education*, 7(2), 133–144.
- Reeve, J. (2009). Why teachers adopt a controlling motivating style toward students and how they can become more autonomy supportive. *Educational Psychologist*, 44(3), 159–175.
- Reeve, J., Vansteenkiste, M., Assor, A., Ahmad, I., Cheon, S. H., Jang, H., Kaplan, H., Moss, J. D., Olaussen, B. S., & Wang, C. K. J. (2014). The beliefs that underlie autonomy-supportive and

controlling teaching: A multinational investigation. *Motivation and Emotion*, 38, 93–110.

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Press.

Author Information

Athraa Abd Ali Lateef , (General Directorate of
Education in Wasit Governorate , Ministry of
Education)
athraaalaydie@gmail.com



Introducing Effort Atrophy Theory through a Comparative Study of Iranian and Iraqi EFL High School Students

Zahraa Mohammed Hussein

Article Info

Article History

Received:

1 March 2026

Revised

1 May 2026

Accepted:

1 July 2026

Published

26 July 2026

Keywords

Effort Atrophy Theory (EAT),
unearned success,
grade inflation,
EFL demotivation,
comparative study

Abstract

This research presents Effort Atrophy Theory (EAT), which proposes that repeated exposure to unearned academic success, particularly through grade inflation practices, gradually diminishes student effort in English as a Foreign Language (EFL) contexts. Data were collected using a quantitative comparative survey design through the newly developed Unearned Success Impact Inventory (USII) administered to 120 Iranian and 120 Iraqi high school EFL teachers. Findings suggested that Iranian teachers were more likely than their Iraqi counterparts to have been compelled to inflate grades due to institutional competition pressure. Results indicated that Iranian students, as perceived by their teachers, reduced their effort considerably more after repeated unearned success experiences. This effort-reward imbalance was mediated by an attributional shift from internal controllable factors (e.g., effort) to external uncontrollable factors (teacher leniency). EAT is proposed and represented through four propositions: (1) Effort-Reward Disconnection Principle, (2) Unearned Success Saturation Effect, (3) Attributional Corrosion Mechanism, and (4) Effort-Reward Imbalance Amplifier. These findings have important implications for teacher professional development, institutional accountability, and curriculum design in high-stakes, competition-driven educational institutions. However, conclusions should be interpreted cautiously given the study's reliance on teacher perceptions and cross-sectional design.

1. Introduction

Effort and academic success have long been linked and considered fundamental to educational motivation (Pekrun, 2006). Learner intrinsic and extrinsic motivation persist over time, in line with Deci and Ryan (1985; Ryan & Deci, 2017), because learners perceive a direct contingent connection between what they exert over time and what they achieve (e.g., grade recognition). However, when this contingency is consistently disrupted, for instance, when success is achieved independently of effort, the motivation structure may be undermined, resulting in a phenomenon known as effort atrophy (Sakai & Kikuchi, 2009; Dörnyei, 2001).

A common but under-theorized violation of the effort-outcome contingency occurs through unearned success, which in this case takes the form of grade inflation, giving students higher grades than they actually deserve (Grade Inflation, n.d.; Kohn, 2002). Although grade inflation has been investigated in Western contexts (Bowden, Rodriguez, & Weingarten, 2025; Fordham Institute, 2024), it also affects non-Western, high power-distance education systems such as those in Iran and Iraq, yet remains sorely understudied in these settings.

In these contexts, institutional competition for status often forces administrators to pressure teachers to inflate grades, creating a cycle of unearned success that systematically undermines student effort. Student-reported study engagement, persistence, homework completion, and internal/controllable versus external/uncontrollable attributions are predominant variables in this study. Institutional pressure is investigated as a contextual moderator.

The overall goal of this research is to present and empirically establish Effort Atrophy Theory (EAT), which argues that repeated experiences of undeserved success may produce long-term, potentially irreversible decreases in learner effort through specific psychological and behavioral mechanisms. Toward that end, we conducted a quantitative comparative study of 120 Iranian and 120 Iraqi high school EFL teachers, examining the prevalence and effects of grade inflation practices on student effort.

The significance of this study is threefold. Theoretically, it presents and operationalizes EAT as a fresh framework for understanding how unearned success deters learner motivation. Practically, the findings are actionable for teacher professional development, institutional policy reform, and curricular design in Iran, Iraq, and similar high power-distance, competitive educational systems. By comparing two neighboring countries sharing cultural origins and recent educational backgrounds, the study also identifies context-specific variations in policy-related practices. Finally, the study illuminates how

unearned success is not merely an administrative convenience but a potentially destructive dimension that undermines the effort that educational systems purport to nurture.

2. Literature Review

2.1 Theoretical Background

2.1.1 Self-Determination Theory (SDT): Quest for Competence

SDT (Deci & Ryan, 1985) identifies three basic psychological needs, autonomy, competence, and relatedness, whose satisfaction is necessary for intrinsic motivation. Competence (Ryan & Deci, 2017), defined as the requirement to feel effective in one's environment, is placed at great risk by unearned success. When students gain high grades without having mastered the material, the cycle that facilitates competence development is broken. According to SDT, motivation that is externally regulated (e.g., by grades) may weaken intrinsic interest when it is perceived as controlling rather than informational (Niemic & Ryan, 2009). Unearned grades are particularly controlling and non-informational, thus eroding the very competence feedback that sustains effort.

2.1.2 Attribution Theory (AT)

Weiner (1985, 1992) introduced AT to explain how people perceive success and failure along three dimensions: locus (internal/external), stability, and controllability. Attributing outcomes to effort, which is internal and controllable, favors persistence following failure (because effort can be improved). Attributing outcomes to ability, internal but stable and uncontrollable, can result in hopelessness. External attributions (e.g., teacher leniency, luck) are uncontrollable and diminish effort, as outcomes are perceived to depend on external factors rather than personal investment. When success is achieved without effort and this pattern is repeated, students systematically transition from internal controllable attributions (effort) to external uncontrollable attributions ("teachers give high grades regardless"), which we predict will diminish future effort (Ahmed et al., 2013; Weiner, 1992).

2.1.3 Expectancy-Value Theory (EVT)

Eccles and Wigfield (2002, 2020) suggest that motivation is a function of (a) expectancy of success and (b) subjective task value. Grade inflation artificially raises expectations to near certainty, massively reducing the perceived "cost" of effort: if success is guaranteed without studying, effort becomes economically irrational. Additionally, unearned success reduces the attainment value of grades (how valuable it is to do well) and can diminish intrinsic interest value, whereby learning becomes separated from meaningful feedback (Wigfield & Eccles, 2000).

2.1.4 Learned Helplessness (LH)

LH theory (Seligman, 1975) describes a state in which recurrent negative events from uncontrollable sources lead to reduced effort to affect outcomes. Classical LH focuses on aversive outcomes. However, we propose a parallel “unearned helplessness” that occurs when positive outcomes (success) become uncontrollable (i.e., can happen regardless of effort). Student disengagement from the learning process can be even more pronounced in such cases (Ferris-Crisafulli, 2022). This is the essence of EAT: unearned success creates a positive-conditioned form of learned helplessness.

2.1.5 Learned Industriousness

According to Eisenberger (1992), the learned industriousness principle holds that effort is rewarded when it is continually reinforced, and its secondary reward characteristics become associated with increased future effort. Unearned success systematically opposes this process: when low effort yields the same high reward, the “effort → reward” link is severed, and learners become “industriously lazy,” learning that minimal effort is sufficient.

2.1.6 Effort-Reward Imbalance (ERI)

The ERI model (Siegrist, 1996), initially proposed for occupational health, asserts that discrepancies between effort and reward cause stress. In education, a different imbalance occurs: low effort coupled with high reward (unearned success). This paradoxical “effort-reward surplus” is not benign; it has been found to predict lower learning engagement and to mediate learned helplessness (Frontiers in Psychology, 2024). We introduce ERI as an important amplifier in EAT.

2.1.7 Social Contagion of Demotivation

Frenzel et al. (2009) found that teacher emotions contagiously affect students (and vice versa). Classroom practices, particularly grade inflation when institutionalized, create a demotivational climate that spreads through the classroom, normalizing low effort and eroding collective academic engagement. This social contagion mechanism is especially apparent in high power-distance cultures, where teacher authority is rarely questioned and students may perceive inflated grades as normative expectations rather than anomalies.

2.2 Prior Related Work

Numerous studies have shown that grade inflation diminishes student effort and achievement. According to the Fordham Institute (2024), when students receive higher grades without commensurate learning, grades lose meaning and the incentive to earn them disappears. Bowden, Rodriguez, and Weingarten

(2025) found that relaxing high school grading standards mechanically increased GPA but decreased effort and did not improve learning. Similarly, effort-reward imbalance consistently shows a negative relationship with learning engagement, mediated by learned helplessness (Frontiers in Psychology, 2024).

In the Iranian EFL context, studies on demotivation have identified teacher-related determinants such as lack of encouragement and unclear instruction as leading to demotivation (Sakai & Kikuchi, 2009; Gao & Liu, 2022). Iranian EFL students often perceive their success as resulting from external influences like teacher favoritism or luck rather than personal effort (Jahedizadeh et al., 2016).

Comparative studies of formal education in Iran and Iraq (Educational Leadership Research, 2025) have shown differences in educational system structure, curriculum content, and the role of government. However, no comparative study has systematically tested grading practices or the consequences of unearned success between these two countries.

Importantly, a unified theory of effort atrophy resulting from unearned academic success in EFL contexts, particularly in a comparative Iran-Iraq setting, has not been proposed or empirically tested. This gap is significant, as unearned success may be more detrimental to long-term learner motivation than commonly acknowledged, and its mechanisms may be contingent on cultural and institutional factors. Despite decades of attention to motivation, demotivation, and grade inflation, three gaps remain:

1. There is no comprehensive theory explaining how repeated unearned success leads to the steady erosion of student effort. Existing theories (SDT, AT, EVT, LH) explain components, but none integrate them into a holistic model of effort atrophy as a clear motivational process.
2. Comparative research on grade inflation practices and motivation-related effects in non-Western, high power-distance educational contexts is virtually nonexistent. The distinctive dynamics of institutional competition in Iran have produced a unique ecology of unearned success that has not been studied empirically.
3. There is no comprehensive examination of the Iranian versus Iraqi EFL context regarding grade inflation frequency, attributional shifts, and effort reduction.

To address these gaps, this study aims to answer the following research questions:

RQ1: To what extent do Iranian and Iraqi EFL high school teachers report engaging in or being pressured to engage in grade inflation practices?

RQ2: Is there a significant difference in self-reported effort reduction among Iranian and Iraqi EFL students, as perceived by their teachers, after repeated exposure to unearned success?

RQ3: Are attributional patterns (internal-controllable versus external-uncontrollable) related to the relationship between unearned success and student effort reduction?

RQ4: Based on the data, what constitute the core components of EAT?

3. Methodology

3.1 Research Design

This study employed a quantitative comparative survey design. This design was chosen because it allowed for the collection of standardized, comparable data across representative samples from two countries and facilitated statistical comparison of grade inflation practices and their effects. The independent variable was the frequency and intensity of unearned success (operationalized as grade inflation practices), while the dependent variables included student effort reduction, motivational engagement, and attributional patterns. Institutional pressure (administrative control by school administration) was considered a contextual moderator.

The comparative survey design was selected for its ability to efficiently capture data from a large sample across two distinct educational systems, enabling cross-national comparisons while maintaining methodological consistency. However, as a cross-sectional design, it cannot establish causality; therefore, results should be interpreted as associations rather than causal relationships.

3.2 Participants

Participants were 240 high school EFL teachers (120 Iranian, 120 Iraqi), selected through stratified random sampling. Stratification was based on school type (public vs. private) and geographic region (urban vs. rural). Participants were recruited based on the following criteria: (1) having taught for at least 3 years at the high school level, (2) currently teaching EFL, and (3) having knowledge of their school's grading policies. The mean teaching experience was 8.7 years ($SD = 4.2$). Gender distribution was approximately equal (52% female, 48% male). The Iranian sample included four provinces (Tehran, Isfahan, Khorasan Razavi, and Fars); the Iraqi sample included three governorates (Baghdad, Basra, and Erbil). Informed consent was obtained from all participants, and ethical approval was obtained from the researchers' institutions. While stratified random sampling enhances representativeness, the sample may not fully reflect all regions within each country. Additionally, the reliance on teacher volunteers may introduce selection bias.

3.3 Research Instrument

The newly developed Unearned Success Impact Inventory (USII) consists of 32 items across four subscales:

1. Frequency of Grade Inflation (8 items): e.g., “How often are you pressured by school administration to give higher grades than students actually merit?”; “How often do you grade effort rather than actual mastery?”
2. Institutional Pressure Intensity (8 items): e.g., “My school expects teachers to maintain high average grades to attract more students.”
3. Student Effort Reduction (8 items): e.g., “Students in my class show less study effort after receiving unexpectedly high grades.”
4. Attributional Pattern (8 items): e.g., “My students attribute their success to my leniency rather than their own study.”

Responses were rated using a 5-point Likert scale (1 = never/strongly disagree; 5 = always/strongly agree). The USII was prepared in English and subsequently translated into Persian and Arabic through a blind back-translation process. The questionnaire was pilot-tested on 30 teachers (15 per country), yielding Cronbach’s alpha values from 0.84 to 0.92 for the subscales. The USII relies entirely on teacher perceptions of student effort and attributions. While teachers are well-positioned to observe behavioral changes, the lack of direct student data is a significant limitation. Future studies should triangulate teacher reports with student self-reports and behavioral measures.

3.4 Data Collection

Data were collected over three months (September to November 2025). In both countries, local assistant teachers distributed the USII in paper format during monthly meetings. Participation was voluntary and anonymous, with teachers returning completed questionnaires in sealed envelopes. Two reminder visits were conducted at each school to improve response rates. In total, 260 questionnaires were administered, with 240 valid responses returned (92.3% response rate). Data entry was subjected to double-checking for accuracy.

3.5 Data Analysis

Data were analyzed using IBM SPSS v.29. Descriptive statistics and normality tests (Shapiro-Wilk) were conducted, along with reliability analysis (Cronbach’s alpha). For RQ1, independent samples t-

tests were used to compare Iranian and Iraqi teachers on grade inflation frequency and institutional pressure. For RQ2, ANCOVA was conducted to compare student effort reduction between the two groups, with years of teaching experience as a covariate. For RQ3, mediation analysis (PROCESS macro v.4.1) was used to investigate whether attributional patterns mediated the relationship between unearned success and effort reduction. For RQ4, qualitative thematic analysis of open-ended responses (n = 78) was performed to refine EAT propositions. For all inferential tests, the alpha level was set at $p < 0.05$. The qualitative thematic analysis was conducted on open-ended responses provided by teachers at the end of the USII questionnaire. These responses were independently coded by two researchers, with disagreements resolved through consensus. Themes were identified through an iterative process of coding, categorization, and synthesis. This approach was intended to complement quantitative findings and to provide richer contextual insights for theory development. However, the qualitative component is exploratory rather than confirmatory, and the propositions should be viewed as preliminary. Assumptions of normality, homogeneity of variance, and linearity were tested and met for all quantitative analyses. Bootstrapping (5,000 resamples) was used in the mediation analysis to ensure robust confidence intervals.

4. Results

4.1 Preliminary Analysis

The Shapiro-Wilk test indicated that all variables were normally distributed ($p > 0.05$). Cronbach's alpha for the total USII was 0.91 (Iran: 0.89; Iraq: 0.90). No outliers were identified.

Table 1. Descriptive Statistics for Key Variables (N = 240)

Variable	Iran (n=120)	Iraq (n=120)	Total (N=240)
Frequency of Grade Inflation (M, SD)	4.12 (0.67)	3.05 (0.89)	3.58 (0.92)
Institutional Pressure (M, SD)	4.31 (0.59)	2.94 (0.97)	3.62 (1.01)
Student Effort Reduction (M, SD)	3.98 (0.72)	2.87 (0.83)	3.42 (0.98)
External Attribution (M, SD)	3.85 (0.81)	2.74 (0.91)	3.29 (0.99)

Note: Scores range from 1 to 5. Higher scores indicate higher frequency/intensity.

4.2 Results for Research Question 1

To address RQ1, the extent to which Iranian and Iraqi high school EFL teachers reported engaging in or being pressured to engage in grade inflation practices, independent samples t-tests were conducted. Results showed statistically significant differences, with Iranian teachers reporting significantly more frequent grade inflation. These findings are detailed in Table 2.

Table 2. Independent Samples t-Test Results for Grade Inflation Frequency

Group	N	Mean (SD)	t	df	p	Cohen's d	95% CI Mean Difference
Iran	120	4.12 (0.67)	10.72	238	<0.001**	1.39	[0.87, 1.27]
Iraq	120	3.05 (0.89)					

**p < 0.01 (two-tailed). Effect size is large.

Similarly, institutional pressure scores varied substantially between the two groups (Table 3).

Table 3. Independent Samples t-Test Results for Institutional Pressure

Group	N	Mean (SD)	t	df	p	Cohen's d	95% CI Mean Difference
Iran	120	4.31 (0.59)	13.41	238	<0.001**	1.73	[1.17, 1.57]
Iraq	120	2.94 (0.97)					

**p < 0.01 (two-tailed). Effect size is large.

Iranian EFL teachers reported significantly higher rates of mandated grade inflation and stronger perceptions of institutional pressure compared to Iraqi teachers. The large effect sizes ($d > 1.39$) indicate that these differences are practically meaningful. Specifically, 82% of Iranian teachers reported being forced by school administration to inflate grades at least weekly, compared to only 17% of Iraqi teachers. These findings reflect teacher perceptions, which may be subject to social desirability bias, particularly in reporting sensitive practices such as grade inflation. Additionally, cross-national differences in survey response styles (e.g., tendency to use extreme response options) may influence comparisons.

4.3 Results for Research Question 2

To address RQ2, whether there was a significant difference in self-reported effort reduction between Iranian and Iraqi EFL students, as perceived by teachers, ANCOVA was conducted with country as the independent variable, student effort reduction as the dependent variable, and teachers' years of experience as a covariate. The covariate was not significant ($p = 0.31$) and was therefore removed from the final model. Results are shown in Table 4.

Table 4. ANCOVA Summary for Student Effort Reduction by Country

Source	SS	df	MS	F	p	η^2 partial	95% CI
Country	74.89	1	74.89	112.40	<0.001**	0.32	[0.24, 0.40]
Error	158.60	238	0.67				
Total	233.49	239					

**p < 0.01. η^2 partial = 0.32 suggests a sizable effect.

Table 5 presents the estimated marginal means.

Table 5. Estimated Marginal Means for Student Effort Reduction

Country	Mean (SE)	95% Confidence Interval
Iran	3.98 (0.07)	[3.84, 4.12]
Iraq	2.87 (0.08)	[2.71, 3.03]

After controlling for the non-significant effect of teaching experience, Iranian students showed statistically significantly greater effort reduction than Iraqi students, as perceived by their teachers. The large effect size ($\eta^2 = 0.32$) indicates that approximately one-third of the variance in effort reduction is explained by country membership. This finding supports the basic premise of EAT: repeated unearned success begets progressive effort atrophy. These findings are based on teacher perceptions of student effort reduction, not direct measures of student behavior. Teachers may overestimate or underestimate changes in student effort, and their perceptions may be influenced by their own beliefs and expectations. Future research should include student self-reports, behavioral observations, and objective measures of effort (e.g., homework completion rates, time spent studying) to validate these findings.

4.4 Results for Research Question 3

To address RQ3, whether attributional patterns (internal controllable vs. external uncontrollable) mediate the relationship between unearned success and student effort reduction, mediation analysis was conducted using PROCESS Model 4 with 5,000 bootstrap samples. Unearned success (grade inflation frequency) was treated as the independent variable, student effort reduction as the dependent variable, and external attribution as the mediator. Results are summarized in Tables 6 and 7.

Table 6. Impact of Unearned Success on Effort Reduction: Direct and Indirect Effects

Type of Effect	Estimate	SE	95% Bootstrap CI	p	Proportion Mediated
Direct effect (c')	0.31	0.06	[0.19, 0.43]	<0.001**	59%
Indirect effect (a × b)	0.24	0.05	[0.14, 0.34]	<0.001**	41%
Total effect (c)	0.55	0.08	[0.39, 0.71]	<0.001**	100%

$p < 0.01$ (using bias-adjusted bootstrap confidence intervals not including zero).

Table 7. Path Coefficients for Mediation Model

Path	Coefficient	SE	t	p	95% CI
a (UE → EA)	0.62	0.07	8.86	<0.001**	[0.48, 0.76]
b (EA → ER)	0.39	0.06	6.50	<0.001**	[0.27, 0.51]
c' (UE → ER)	0.31	0.06	5.17	<0.001**	[0.19, 0.43]

Note: UE = Unearned Success (frequency of grade inflation); EA = External Attribution; ER = Effort Reduction. All coefficients are unstandardized.

External attribution significantly mediated the association between unearned success and effort reduction. The indirect effect (0.24) was statistically significant, as the bootstrap confidence interval did not include zero. The partial attributional shift from internal controllable (effort) to external uncontrollable (teacher factors) accounted for approximately 41% of the total effect of unearned success on effort reduction. The direct effect remained significant, suggesting partial mediation. The mediation model assumes the correct temporal ordering of variables (unearned success → external attribution → effort reduction), but the cross-sectional design cannot establish causality. Alternative pathways (e.g., effort reduction leading to increased external attribution) cannot be ruled out. Longitudinal research is needed to confirm the causal direction.

4.5 Results for Research Question 4

To address RQ4, identifying the core components of EAT as derived from the data, qualitative thematic analysis of open-ended responses ($n = 78$) was conducted. The analysis was guided by Braun and Clarke's (2006) six-phase framework: familiarization with data, initial coding, theme generation, theme review, theme definition, and write-up. Two independent coders analyzed the responses, with inter-coder reliability of 0.87. The analysis produced four key propositions that, together with the quantitative evidence, constitute EAT.

Proposition 1: Effort-Reward Disconnection Principle

The cognitive link between effort and reward is impaired by repeated instances of success that are not associated with effort. Teachers noted that students “expect high grades at any cost” and feel “they do not need to study anymore.” This disconnection was particularly pronounced in the Iranian sample, where teachers reported that students had come to view high grades as an entitlement rather than an earned outcome. This proposition integrates SDT's competence need frustration (Ryan & Deci, 2017) with EVT's devaluation of effort (Eccles & Wigfield, 2002), suggesting that repeated unearned success disrupts the fundamental contingency between effort and outcome that sustains motivation.

Proposition 2: Dose-Response Effect of Unearned Success Saturation

The effect on effort is dose-dependent: each episode of unearned success incrementally weakens the effort-outcome contingency, with cumulative effects becoming substantial over time. Iranian teachers noted that after the 3rd or 4th unearned high grade, “students no longer even tried.” This suggests a saturation effect, where initial episodes may have minimal impact, but repeated experiences lead to

progressive erosion of effort. This proposition aligns with the learned helplessness framework (Seligman, 1975; Ferris-Crisafulli, 2022), extending it to positive but uncontrollable outcomes. It also resonates with the learned industriousness principle (Eisenberger, 1992), where the reinforcement of effort is systematically undermined.

Proposition 3: Attributional Corrosion Mechanism

Unearned success shifts student attributions from internal controllable factors (effort) to external uncontrollable factors (teacher leniency, luck, ease of grading). This shift occurs through an organic series of steps: initial confusion, gradual acceptance, and eventual expectation. As one Iranian teacher described, “Students ask themselves: ‘Why should I study? The teacher will let me pass anyway.’” This proposition directly draws on Weiner’s (1985, 1992) attribution theory, providing empirical support for the prediction that external, uncontrollable attributions diminish future effort. It also extends the theory by specifying the mechanism through which repeated unearned success leads to attributional change.

Proposition 4: Effort-Reward Imbalance Amplifier

The mismatch between minimal effort and maximal reward creates a motivational “surplus” that paradoxically decreases future effort investment. This imbalance is not benign; it amplifies the other mechanisms, making effort atrophy more likely. Teachers reported that when students perceive that “no effort is needed for good grades,” the reward system loses its incentive value entirely. This proposition integrates the effort-reward imbalance model (Siegrist, 1996; *Frontiers in Psychology*, 2024) and extends it to educational contexts. It suggests that the paradoxical “effort-reward surplus” is a key amplifier of the other EAT mechanisms. The propositions derived from open-ended responses are exploratory and should be viewed as preliminary. The qualitative data provide rich illustrations of the quantitative patterns, but further empirical validation, through longitudinal studies, student data, and experimental designs, is necessary to confirm these propositions. The propositions are presented as a theoretical framework to guide future research, not as established facts.

5. Discussion

5.1 Discussion of Research Question 1

The first research question assessed the extent to which Iranian and Iraqi EFL high school teachers engaged in or were pressured to engage in grade inflation. The findings demonstrated that Iranian EFL teachers reported significantly higher frequency of grade inflation and institutional pressure than Iraqi teachers, with large effect sizes (Cohen’s $d = 1.39$ and 1.73 , respectively).

In SDT (Deci & Ryan, 1985; Ryan & Deci, 2017), competence feedback is considered crucial to the development of intrinsic motivation. When teachers are pressured to boost grades, they provide false competence feedback that violates students' need for honest self-assessment and teachers' own professional ethical duties. In high power-distance societies like Iran, hierarchical pressure from administrators tends to substitute for teacher autonomy, resulting in a systemic demotivational ecology.

According to AT (Weiner, 1985, 1992), learners exposed to such systems quickly learn to conflate success with external factors (teacher decisions) rather than attributing achievement to internal factors. These results are consistent with findings by Bowden, Rodriguez, and Weingarten (2025), who reported that lowering high school grading standards increased GPA but decreased student effort. The Fordham Institute (2024) found similar results: lenient grading hurts student outcomes. Our research extends these findings by identifying sharp cross-national differences between Iran and Iraq and linking these differences to varying levels of institutional competition pressure. The observed differences may also reflect reporting biases. Iranian teachers may be more willing to report grade inflation due to greater awareness of the issue, whereas Iraqi teachers may underreport due to social desirability concerns. Cross-cultural differences in response styles may also contribute to the observed differences.

5.2 Discussion of Research Question 2

The second research question examined whether there was a significant difference in student effort reduction between Iranian and Iraqi EFL students, as perceived by teachers, after repeated exposure to unearned success. The results demonstrated that Iranian students showed significantly greater decreases in effort compared to Iraqi students (large effect size $\eta^2 = 0.32$).

EVT (Eccles & Wigfield, 2002, 2020) suggests a dramatic reduction in the subjective value of effort when success is guaranteed independently of effort. In Iran, where unearned success is widespread, effort becomes economically irrational. This is strengthened by the effort-reward imbalance model (Siegrist, 1996; *Frontiers in Psychology*, 2024), in which the mismatch between minimum effort and maximum reward creates a motivational “surplus” that paradoxically decreases future effort investment, a core effect of EAT.

These findings are consistent with work in the effort-reward imbalance literature demonstrating that low effort combined with high reward is a consistent predictor of poor learning engagement, mediated by learned helplessness (*Frontiers in Psychology*, 2024). It also aligns with the “Dr. Fox” effect (Ware & Williams, 1975), where students exposed to charismatic but content-poor lectures rated them highly but learned little, and this undeserved positive evaluation quieted effort. The measure of student effort

reduction was based on teacher perceptions, not direct student reports. Teachers may have different thresholds for perceiving effort reduction, and their perceptions may be influenced by their own expectations or biases. Additionally, the cross-sectional design means we cannot rule out the possibility that students with lower effort levels are more likely to be in environments with grade inflation, rather than grade inflation causing effort reduction.

5.3 Discussion of Research Question 3

The third research question examined whether attributional patterns (internal controllable vs. external uncontrollable) mediated the association between unearned success and student effort reduction. The results showed that external attribution significantly mediated this association, explaining 41% of the effect.

AT (Weiner, 1985, 1992) provides a direct explanatory mechanism. Repeated unearned success provides evidence that effort is not the origin of success. When students experience undeserved success, they begin to transition from internal controllable attributions (effort) to external uncontrollable attributions (teacher leniency). This shift follows the pattern that attribution theory predicts will reduce future effort, as people are less inclined to invest effort in outcomes they believe they cannot control. This mediation model is consistent with Ahmed et al. (2013), who reported that external attribution of success led to lower subsequent effort. It is also consistent with the extensive literature on learned helplessness (Seligman, 1975; Ferris-Crisafulli, 2022), wherein repeatedly uncontrollable positive events lead to passivity, just as uncontrollable negative events do.

The cross-sectional nature of the data limits causal inference. The temporal ordering assumed in the mediation model (unearned success → external attribution → effort reduction) cannot be confirmed. It is possible that students who are less motivated to begin with are more likely to develop external attributions, which then make them more susceptible to the effects of unearned success. Additionally, other mediators not included in the model (e.g., self-efficacy, goal orientation) may be at play.

5.4 Discussion of Research Question 4

The fourth research question focused on identifying the core components of EAT based on the data. The four data-derived propositions together represent EAT. EAT draws on SDT (competence need frustration), attribution theory (effort → ability/luck shift), EVT (effort devaluation), LH (learned helplessness orientation), LI (extinction of industriousness), and ERI (imbalance amplification).

What makes EAT unique is its specific emphasis on unearned success as the primary cause and effort atrophy as the unique outcome. Although learned helplessness is usually oriented toward aversive

uncontrollability, EAT presents a parallel “unearned helplessness” that arises from non-aversive but uncontrollable positive outcomes. EAT offers a coherent paradigm for understanding the motivational harm resulting from grade inflation, with real implications for educational policy and teacher training in high power-distance, competition-driven environments.

However, the qualitative findings that form the basis of the propositions are exploratory and require further empirical validation. The sample of 78 open-ended responses is relatively small, and the responses were from teachers only. The propositions would be strengthened by triangulation with student interviews and larger qualitative samples.

6. Conclusion

This study developed and provided empirical evidence for EAT through a quantitative comparison of 240 Iranian and Iraqi EFL high school teachers. Key conclusions are as follows:

1. Iranian teachers reported significantly higher frequencies of forced grade inflation and institutional pressure compared to Iraqi teachers, likely due to more intense inter-school competition in Iran. However, these findings reflect teacher perceptions and may be influenced by reporting biases.
2. Iranian students, as perceived by their teachers, showed significantly greater decreases in effort after repeated unearned success, corroborating the EAT premise of progressive effort atrophy following unearned success. These findings are based on teacher reports and should be validated with student-level data.
3. An attributional shift from internal controllable (effort) to external uncontrollable (teacher factors) significantly mediated this relationship, accounting for approximately 41% of the effect. However, the cross-sectional design limits causal inference.
4. Effort Atrophy Theory consists of four interrelated propositions, Effort-Reward Disconnection Principle, Unearned Success Saturation Effect, Attributional Corrosion Mechanism, and Effort-Reward Imbalance Amplifier that together provide a coherent account of why unearned success demotivates learners. These propositions are preliminary and require further empirical validation.

This study has several implications. Existing motivation theories (SDT, AT, EVT, LH, LI, ERI) are further strengthened by EAT, which is based on the concept that unearned success can have unique motivational implications. EAT provides a testable model for future investigations of grade inflation and

academic demotivation. However, the theory should be viewed as a framework for generating hypotheses rather than a confirmed model.

Pre-service and in-service training should recognize the discouraging influence of unearned success and guide teachers on how to counter institutional pressure to inflate grades. Professional development should focus on accurate, mastery-based assessment strategies that differentiate between effort, behavior, and academic achievement. Teachers should be equipped with skills to provide honest, constructive feedback that supports student motivation.

Schools should be aware that competition centered on inflated grade averages has the opposite effect, fostering a culture of unearned success that diminishes true student effort and learning. Policies that encourage the use of visible, externally validated test measures, such as standardized benchmarks, can help mitigate grade inflation. Accountability systems should reward genuine student achievement rather than artificially inflated grades.

Multiple authentic, low-stakes assessments that provide valid, correct feedback on competence should be incorporated into curricula. Assessment should not be reduced to a single point, the final grade, but should provide ongoing feedback that supports learning. Constructive feedback that focuses on effort and improvement can help maintain the effort-outcome contingency. In addition, several limitations should be noted:

This study relied entirely on teacher self-reports. Teachers' perceptions of student effort reduction and attributional patterns may not correspond to students' actual experiences, behaviors, or motivational states. Social desirability bias may also affect reporting, particularly regarding sensitive practices like grade inflation. Future research should include student self-reports, behavioral measures, and observational data to triangulate findings.

The cross-sectional design inhibits causal inference. While the mediation model assumes a specific causal sequence (unearned success → attribution → effort reduction), alternative pathways cannot be ruled out. Longitudinal studies are needed to establish directionality and to examine how the effects of unearned success unfold over time.

Although stratified random sampling was used, the sample may not fully represent all regions within Iran and Iraq. The samples were drawn from specific provinces and governorates, and findings may not generalize to rural areas or to regions with different educational policies. The sample size, while adequate for the quantitative analyses, may be insufficient for detecting smaller effects or for subgroup analyses.

The instruments were translated and back-translated to ensure linguistic equivalence, but cultural differences in response styles (e.g., tendencies toward extreme responses, acquiescence) may affect comparisons. Additionally, the constructs being measured, such as “effort” and “grade inflation,” may have different meanings in the Iranian and Iraqi educational contexts.

The proposed EAT is based on limited evidence from a single cross-sectional survey. The qualitative propositions, while theoretically grounded, are exploratory and require stronger empirical validation. The theory should be tested through multiple methods, including experiments, longitudinal designs, and studies in different educational contexts.

The absence of direct student data is a significant limitation. Teachers’ perceptions may be influenced by their own experiences, biases, and expectations. Students’ actual motivational states, effort levels, and attributional patterns may differ from teachers’ perceptions. Future research should collect data from both teachers and students to examine convergence and divergence in perspectives.

The findings are specific to EFL contexts in Iran and Iraq and may not generalize to other subjects, other educational levels, or other countries. Cultural factors, institutional structures, and educational policies vary widely, and the mechanisms proposed by EAT may operate differently in different contexts. Building on the current findings, future research should address the following:

1. Conduct longitudinal studies to examine the dose-response relationship between unearned success and effort atrophy over time. Such studies could establish the temporal ordering of variables and examine how the effects accumulate.
2. Integrate student surveys directly into investigations to measure students’ attributional patterns, effort levels, and motivational states. Triangulating teacher and student data would provide a more comprehensive picture of the phenomena.
3. Examine differences between public and private schools, as institutional pressures may differ. Private schools may face different market pressures, while public schools may be subject to different accountability systems.
4. Test interventions designed to restore the effort-outcome contingency. For example, mastery-based grading systems, transparent assessment criteria, and constructive feedback practices could be evaluated for their effectiveness in mitigating effort atrophy.

5. Extend the application of EAT to other high-stakes or competition-driven educational contexts beyond Iran and Iraq, such as Turkey, Pakistan, Egypt, and other countries with similar educational dynamics.
6. Conduct experimental studies that manipulate the presence of unearned success to establish causality. Random assignment to conditions with or without grade inflation would provide stronger evidence for the causal claims of EAT.
7. Use qualitative methods to capture students' lived experiences of unearned success and its motivational consequences. In-depth interviews, focus groups, and diary studies could provide richer insights into the psychological mechanisms at play.
8. Investigate moderators of the relationship between unearned success and effort reduction, such as individual differences in achievement goals, self-efficacy, and personality characteristics.
9. Explore the role of teacher beliefs and attitudes in the perpetuation of grade inflation. Understanding why teachers comply with or resist institutional pressure could inform professional development interventions.
10. Examine the long-term consequences of effort atrophy, including its effects on academic achievement, career trajectories, and lifelong learning outcomes.

7. References

- Ahmed, W., van der Werf, G., Kuyper, H., & Minnaert, A. (2013). Emotions, self-regulated learning, and achievement in mathematics: A growth curve analysis. *Journal of Educational Psychology*, 105(1), 150–161.
- Bowden, M., Rodriguez, J., & Weingarten, R. (2025). How does academic leniency affect student effort, achievement, and long-run human capital? *The Policy Scientist*.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum Press.
- Dörnyei, Z. (2001). *Motivational strategies in the language classroom*. Cambridge University Press.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53, 109–132.

- Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, 61, 101859.
- Educational Leadership Research. (2025). Comparative study of formal education in Iran and Iraq. *Educational Leadership Research Journal*.
- Eisenberger, R. (1992). Learned industriousness. *Psychological Review*, 99(2), 248–267.
- Ferris-Crisafulli, J. (2022). "Even the little things matter:" A phenomenological study on factors impacting student motivation during and after COVID-related disruptions in education (Doctoral dissertation). University of South Carolina.
- Fordham Institute. (2024). Grade inflation: Why it matters and how to stop it. *The Education Gadfly Weekly*.
- Frenzel, A. C., Goetz, T., Lüdtke, O., Pekrun, R., & Sutton, R. E. (2009). Emotional transmission in the classroom: Exploring the relationship between teacher and student enjoyment. *Journal of Educational Psychology*, 101(3), 705–716.
- Frontiers in Psychology. (2024). Effort-reward imbalance and learning engagement: A meta-analytic review. *Frontiers in Psychology*, 15.
- Gao, L., & Liu, H. (2022). Revisiting students' foreign language learning demotivation: From concepts to themes. *Frontiers in Psychology*, 13, 1030634.
- Grade inflation. (n.d.). In *Oxford English Dictionary*.
- Jahedizadeh, S., Ghanizadeh, A., & Ghonsooly, B. (2016). The role of EFL teachers' emotional intelligence in their success. *Iranian Journal of Language Teaching Research*, 4(2), 65–83.
- Kohn, A. (2002). *The case against standardized testing: Raising the scores, ruining the schools*. Heinemann.
- Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and Research in Education*, 7(2), 133–144.
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18(4), 315–341.
- Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Press.

- Sakai, H., & Kikuchi, K. (2009). An analysis of demotivators in the EFL classroom. *System*, 37(1), 57–69.
- Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death*. W. H. Freeman.
- Siegrist, J. (1996). Adverse health effects of high-effort/low-reward conditions. *Journal of Occupational Health Psychology*, 1(1), 27–41.
- Ware, J. E., & Williams, R. G. (1975). The Dr. Fox effect: A study of lecturer effectiveness and ratings of instruction. *Journal of Medical Education*, 50(2), 149–156.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92(4), 548–573.
- Weiner, B. (1992). *Human motivation: Metaphors, theories, and research*. Sage.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68–81.

Author Information

**Zahraa Mohammed Hussein (General Directorate
of Education in Wasit Governorate , Ministry of
Education)**
zah84ra@gmail.com



The Effect of Multimodal Instruction on Vocabulary Learning and Retention in Iraqi EFL Classrooms

Rusul Esmaeel Sakran

Article Info

Article History

Received:

1 March 2026

Revised

1 May 2026

Accepted:

1 July 2026

Published

27 July 2026

Keywords

Multimodal instruction, vocabulary learning, vocabulary retention, Iraqi EFL learners, multimedia learning

Abstract

Vocabulary knowledge remains a fundamental component of second language competence, yet Iraqi EFL learners face considerable challenges in acquiring and retaining lexical items due to limited exposure to authentic English and predominant reliance on traditional text-based instruction. This quasi-experimental study examined the impact of multimodal instruction on vocabulary learning and retention among 60 intermediate Iraqi EFL learners randomly assigned to an experimental group ($n = 30$) receiving instruction integrating visual, auditory, and textual materials, or a control group ($n = 30$) taught through conventional text-based methods. While extensive international research has documented the benefits of multimodal approaches for vocabulary acquisition, empirical evidence from the Iraqi EFL context, characterized by resource constraints, examination-driven pedagogy, and limited authentic language exposure, remains conspicuously absent. This study addresses this gap by drawing upon Dual Coding Theory (Paivio, 1971), Cognitive Theory of Multimedia Learning (Mayer, 2014), and Depth Processing Hypothesis (Craik & Lockhart, 1972) to investigate whether multimodal instruction yields superior learning outcomes in this under-researched context. Both groups completed a 30-item vocabulary test before intervention, immediately after, and two weeks later. One-way ANCOVA controlling for pre-test scores revealed that the multimodal group demonstrated significantly better performance on both the immediate post-test ($F(1,57) = 40.62, p < .001, \eta^2 = .413$) and the delayed post-test ($F(1,57) = 62.31, p < .001, \eta^2 = .521$), with large effect sizes indicating substantial practical significance. These findings empirically confirm that the simultaneous presentation of written definitions, visual illustrations, audio recordings, and video clips creates redundant memory traces and facilitates deeper semantic processing, thereby enhancing both initial vocabulary acquisition and long-term retention. The results provide strong theoretical validation for multimodal learning principles in a resource-limited EFL context and offer practical guidance for curriculum developers, material designers, and classroom practitioners seeking evidence-based strategies to improve vocabulary instruction in Iraqi educational settings.

1. Introduction

Vocabulary knowledge is a cornerstone of second-language competence, providing learners with the ability to understand input and to convey meaning in various skill areas (Nation, 2001; Schmitt, 2010). In the absence of a sufficiently large and highly processed vocabulary, learners fail to read authentic texts, follow spoken discourse, and produce coherent speech and writing (Wilkins, 1972). The significance of vocabulary is particularly relevant in English as a Foreign Language (EFL) contexts where exposure to the target language is often only possible in formal or structured formats and where the sheer quantity of words to be learned presents a continuing problem for both learners and teachers (Nation, 2006).

In Iraq, English language teaching has historically been teacher-centred, examination-oriented, and characterized by practices that limit authentic language use outside the classroom, with English often perceived merely as a test subject rather than a communicative tool (Saeed, 2015). The traditional approach, which emphasizes rote memorisation and translation, frequently causes students to struggle with vocabulary acquisition and retention (Al-Marzouqi, 2020). Public schools lack resources such as libraries, audio-visual tools, and additional materials to engage learners effectively, while classrooms suffer from large class sizes, noise pollution, and inadequate infrastructure (Saeed, 2015). These challenges are compounded by an unstable security situation that has resulted in additional interruptions to the educational process and decreased student motivation (Saeed, 2015). Such conditions render traditional text-centred vocabulary instruction largely ineffectual, as it fails to provide learners with the rich contextualised language input necessary for meaningful processing and long-term retention (Al-Marzouqi, 2020).

Given these contextual limitations, educational researchers have increasingly explored multimodal strategies as a means to overcome the shortcomings of traditional teaching by integrating visual, auditory, and textual information to achieve higher learning outcomes. In multimodal instruction, presenting content through various sensory modalities enhances cognitive processing and improves sustained retention (Paivio, 1971; Mayer, 2014). When vocabulary learning incorporates multimodal resources, such as images, video, audio recordings, and written annotations, learners gain access to a broader range of contextual elements and diverse retrieval pathways. This approach not only supports vocabulary acquisition but also promotes long-term retention (Chun & Plass, 1996; Yanguas, 2009). Additionally, it is believed to reduce cognitive load, boost motivation, and foster meaningful engagement with lexical items (Sweller, 1988; Mayer, 2014).

Despite the well-documented merits of multimodal instruction in various EFL settings worldwide, few studies have systematically investigated its impact on vocabulary learning and retention in the Iraqi EFL context. The current research addresses this gap by examining the effect of a comprehensive multimodal instructional package, incorporating written definitions, colour photographs/illustrations, audio recordings of native speaker usage, and contextual video clips, on Iraqi EFL learners' immediate vocabulary learning and their ability to retain acquired vocabulary over time. Using a quasi-experimental design, this study compares the outcomes of students who experienced vocabulary instruction through integrated visual, auditory, and textual materials with those of their counterparts taught using a traditional text-based method. The overarching goal is to investigate whether multimodal instruction leads to greater vocabulary gains and to assess the durability of those gains two weeks after instruction. By grounding this investigation in well-established cognitive theories, Dual Coding Theory, Cognitive Theory of Multimedia Learning, and the Depth Processing Hypothesis, this study aims to provide both theoretical insights and practical guidance for vocabulary instruction in resource-constrained Iraqi EFL classrooms.

2. Literature Review

2.1 Theoretical Background

Multimodal instruction is theoretically based on multiple cognitive and pedagogical theories. Core to these are dual coding theory, cognitive theory of multimedia learning, cognitive load theory, and the depth processing hypothesis.

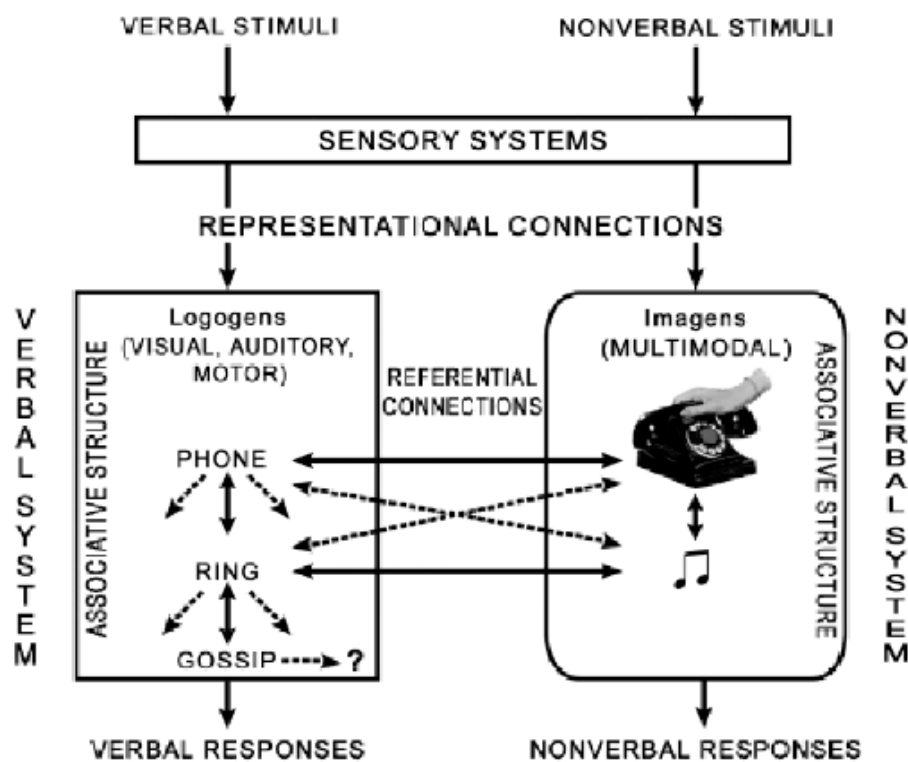
2.1.1 Dual-Coding Theory (DCT)

Allan Paivio's dual-coding theory, put forward in 1971, highlights the impact of mental imagery on memory and cognitive processes. Initially, the theory focused on memory but later expanded to include other cognitive domains (Paivio, 1971, 1986, 1991). The basic claim is that while humans can acquire knowledge through either verbal association or visual imagery, combining the two leads to better learning results (Reed, 2010). DCT suggests that the human brain processes information through two distinct channels: verbal and nonverbal systems. Sternberg (2003) referred to the nonverbal system as the imagery-based system. In the verbal system, linguistic units known as logogens are stored, while the imagery system processes picture images (imagens) (Paivio, 2014).

According to Paivio (1986) and Clark and Paivio (1991), three levels of processing are identified: representational, referential, and associative. Representational processing occurs when the stimulus

directly provokes verbal or visual representation. Referential processing establishes associations between two systems, such as a picture evoking a word or a word calling up an image. This process is crucial for developing language and creative skills (Paivio, 2014). The activation across these systems allows a child to connect a name to an object, gradually leading to mastery of structured speech (Moeser & Bregman, 1973). Associative processing refers to behavior in which an individual draws a connection between units from one system.

Figure 1. *The DCT Model of Multimodal Logogen and Imagen Units, Connections, and Implied Processes* (adopted from Kanellopoulou et al., 2019)



DCT is widely applicable to second language (L2) acquisition. Verbal and nonverbal codes operate independently, creating an additive memory retrieval effect (Paivio & Begg, 1981). When students link a new vocabulary word to both sound and image together, they remember the word better than when using a singular modality (Paivio, 1991). In her 1994 article, Rieber argued that memory task illustrations of new vocabulary work better than text only. A single memory trace may not be enough for effective learning.

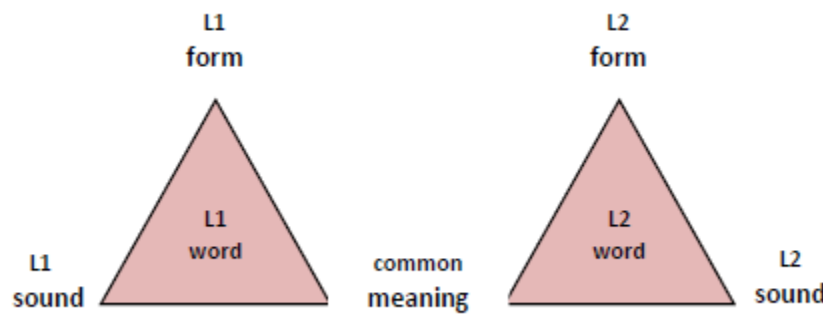
Often, a picture triggers the pathway linking an image to its corresponding word (Lin, 2009), making a second memory trace more accessible. Consequently, a multimedia environment that integrates text, sound, and moving images provides an ideal setting for vocabulary acquisition (Al-Seghayer, 2001, 2016). Nonetheless, DCT suffers a limitation primarily in the “concreteness effect,” where abstract ideas are difficult to visualize and therefore less likely to receive dual coding (Clark & Paivio, 1991). Concrete information is most amenable to processing in both channels, whereas abstract information is generally limited to verbal processing (Paivio, Salinas-Pazajas de la Torre, & Salim, 1994). Still, empirical test data show that DCT increases memory in many cases (Reed, 2010).

In the future (Pylyshyn, 1973), it should also examine if word and image are the sole factors at work in cognitive function or if there are other coding modalities. While Paivio (1971) originally proposed a single language-based model, Paivio and Desrochers (1980) subsequently extended the theory to encompass bilingual or multilingual contexts. In bilingual DCT, there are two separate verbal systems connected to a common imagery system. Scientists using this model have proposed several hypotheses on how the three components, autonomous, interconnected, and interdependent, function together. This study employs dual coding and bilingualism to examine the relationship between words and images, as well as the connection between a person’s first language (L1) and their second language (L2). Soh (2010) outlines several key assumptions that inform these interactions within the context of language learning:

1. As in the original monolingual theory, the imagery system is independent of the two verbal systems.
2. The three systems are linked by referential links so that imagery and verbal systems can affect one another reciprocally.
3. While the two verbal systems have referential connections to the same imagery system, they are also independent and partially related. There are two important implications for this: the extent of overlap will differ with each language pair, and the imagery system serves as a bridge allowing indirect connection between languages.
4. Bilingual people have two separate and yet interrelated verbal systems, so change in language input can lead to code-switching.

Essentially, when students pull up vocabulary from one verbal system and start recoding it in the other verbal system, practice confers on these words fresh meaning. This is a typical way to gain vocabulary with the added skill of generating translation equivalents across languages.

Figure 2. *Sharing Meanings between Two Languages through Code-Switching* (Adopted from Soh, 2010)



Despite significant advancements in this cognitive theory over the years, the bilingual DCT still requires considerable further study to assess its validity. In particular, more research is needed to understand the relationship between a learner’s first language (L1) and second language (L2) and how the verbal systems of these two languages interact. A significant limitation arises when there is a lack of information in the first language (L1) compared to the knowledge in the second language (L2). For example, if a learner tries to understand an L2 concept that lacks an equivalent in L1, then there is no existing connection between the two languages. This lack of connection means that, in such cases, the learner cannot effectively apply the relevant theories. When this link exists, it can either enhance learning or cause interference.

2.1.2 Cognitive Theory of Multimedia Learning (CTML)

CTML studies how people acquire knowledge through a process of multimedia presentation (Mayer, 1997, 2001, 2005, 2014). Drawing on Wittrock’s (1974) generative theory and Paivio’s (1986) DCT, CTML extends the former by framing the learner as an active “knowledge constructor,” not merely a passive recipient of information, who intentionally chooses and associates material, oral, and visual information between two distinct kinds of information (Mayer, 1997). This theory emphasizes that students make relationships between words and images, interpreting images and words in an active process for permanence and memory, so that learning becomes meaningful (Mayer, 1997, 2005). One

major difference between CTML and DCT is that CTML centres around an active, responsible learner in constructing information.

Mayer (2005, 2014) identifies three basic assumptions that underlie the learning process in CTML. The first is based on two cognitive pathways: the auditory/verbal channel and the visual/pictorial channel, like DCT, which postulates that sensory memory processes information through a couple of different pathways: auditory and visual. The first allows for processing auditory input (e.g., narration and spoken word) while the second processes visual input (animation, video, and graph). This type of reasoning rests on the premise that working memory (per cognitive load theory) has a limited capacity to process information in a given channel at one given time. Cognitive overload occurs when task demand exceeds this processing threshold. Mayer (2014) states that working memory is limited to maintaining a small volume of visual or verbal information simultaneously; thus, only a portion of an illustration or narration retained is retained in the relevant mental model at any given time.

Active processing: CTML describes learning as a continually evolving process in which learners choose and construct words and images, then link these with knowledge they have already retrieved from long-term memory. Mayer (2014) explains that integrating new information can involve particular cognitive functions: selecting pertinent verbal and visually available data for working memory, structuring these images into verbal and visual models, and incorporating those models into their own representation from current knowledge.

Multimedia input, particularly words and images, enters sensory memory from the environment through our senses (eyes and ears), as illustrated in Figure 3. Visual sensory memory briefly stores images and written text, while auditory sensory memory temporarily holds speech and sounds. This process occurs when learners focus on selecting particular words and images for further processing. Once this information is selected, it enters working memory, where structural relationships are established, and data is temporarily stored and manipulated.

Mayer (2014) theorizes that working memory consists of two interacting components: raw sensory information (both auditory and visual) and the information generated from it, which consists of verbal and pictorial representations that are interconnected.

Long-term memory serves as a vast storage system for information that can endure over time. Meaningful learning relies on the ongoing interaction between working memory and long-term memory (Mayer, 2005, 2014). In this interaction, relevant prior knowledge stored in long-term memory is

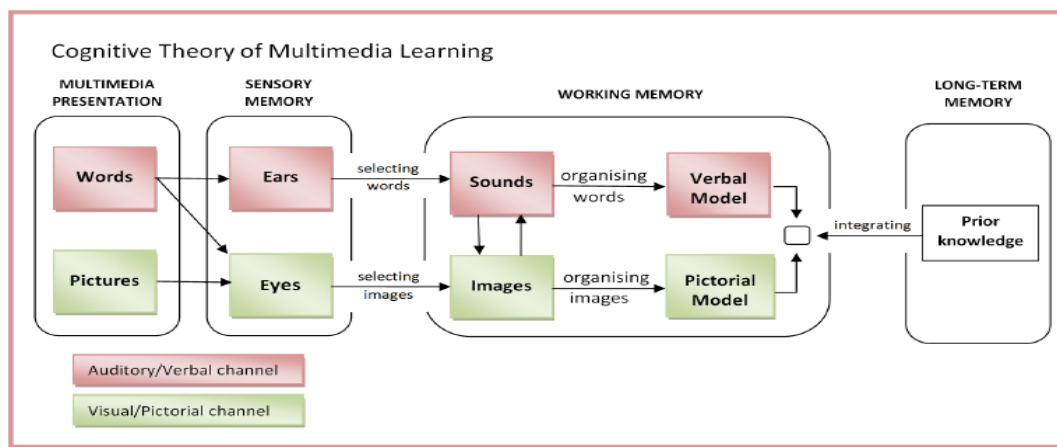
activated and connected with pertinent information entering working memory. This process aids learners in constructing logical mental representations (Mayer, 2014).

CTML follows two core principles: the multimedia principle and the temporal contiguity principle. The multimedia principle aligns with the assumptions behind dual coding theory, which posits that individuals learn more effectively when using a combination of words and images rather than relying on words alone (Mayer, 2014). This principle applies not only to written text and pictures but also to all visual elements, ranging from illustrations to video presentations (Butcher, 2006, 2014), as well as all verbal elements, including text, audio, and sound effects. By presenting information in various formats, we can increase working memory capacity and enhance the ability to recall information from long-term memory later on (Sweller, 2005).

The temporal contiguity principle, on the other hand, mitigates the risk of undue cognitive load by suggesting that deep learning occurs more frequently when corresponding verbal and visual content is presented simultaneously rather than sequentially (Mayer & Fiorella, 2014). Relationships can only be formed between closely related words and images within working memory when these elements are presented at the same time; therefore, simultaneous presentation is usually more effective than separate or consecutive communication (Mayer, 2008; Rusanganwa, 2015).

In addition to these two principles, CTML offers a more comprehensive evidence-informed guide to teaching, which Mayer categorizes based on cognitive functions (Koscianski & Zanutto, 2014). These concepts have been empirically validated (Mayer, 2008) and applied to the design of instructional materials (Herrington & Oliver, 2000; Park & Hannafin, 1993) as well as to predicting learning events (Mayer, 2001).

Figure 3. Cognitive Theory of Multimedia Learning (Adopted from Mayer, 2014, p. 52)



The first group, which is intended to reduce extraneous cognitive load, consists of the coherence effect (removing irrelevant data), signaling effect (focusing on important material to help students track attention), redundancy effect (stopping simultaneous presentation of the same text and narration on screen), spatial contiguity effect (showing similar text and image in the same spot), and temporal contiguity effect (presenting verbal and visual information together). Category two, essential processing, consists of the segmenting effect (segmenting difficult information into clear and digestible steps), pre-training effect (providing important information before the first lesson), and modality effect (using spoken rather than written text when visually presented).

The third class is designed to facilitate generative processing through a multimedia effect that combines text and image, along with a personalization approach that uses a conversational style and addresses the audience directly rather than in a formal third person. In conclusion, two broad frameworks, dual-coding theory (including bilingual dual-coding theory) and the cognitive theory of multimedia learning, reach the conclusion that multimedia conditions enhance learning by simultaneously presenting textual and visual information (Mayer, 2014; Paivio, 1986). Material that links these two modes minimizes mindless mental effort and stimulates more profound processing (Mayer & Fiorella, 2014). Recent insights have increasingly shaped contemporary studies on second language vocabulary acquisition, where watching film in the original language with subtitles is viewed as a highly effective instructional method.

2.1.3 Cognitive Load Theory (CLT)

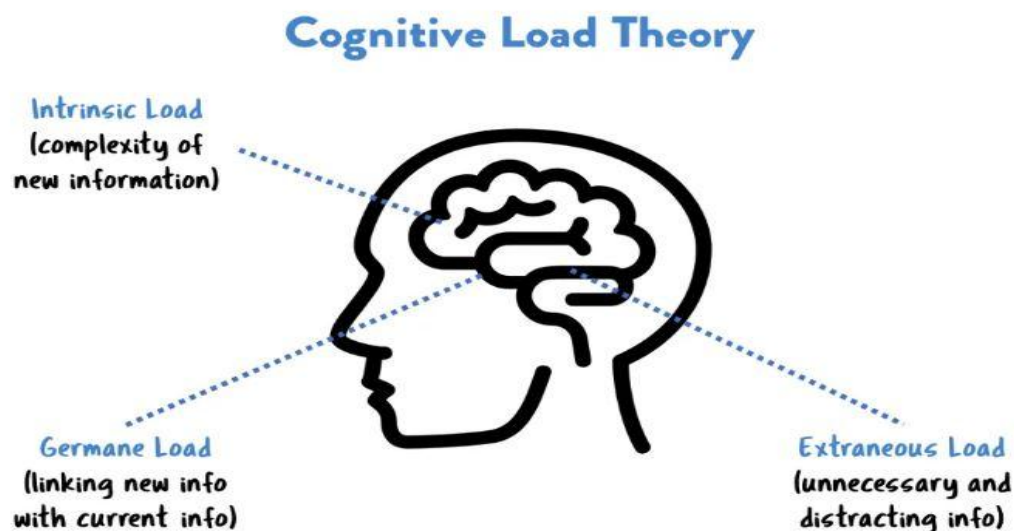
CLT was postulated by cognitive scientists and asserts that human cognitive capacity is essentially constrained. Working memory (WM) is only able to store information that can occur for a small number of items at any one time (Cowan, 2001). While WM has relatively limited capacity to consolidate simple decontextualized knowledge items, long-term memory (LTM) has unlimited capacity to store complex contextualized knowledge structures called “schema.” For CLT, the main focus of instructional design is to modulate cognitive load on WM, which enables schema building in LTM (Paas, Renkl, & Sweller, 2003). This theory is particularly relevant to designing instruction using short videos, as these short films are rich in subject matter and can easily overwhelm learners’ relatively limited WM capacity if not orchestrated properly.

CLT identifies three types of cognitive load that must be managed: intrinsic, extraneous, and germane. Intrinsic cognitive load refers to the inherent difficulty and complexity of the material itself. For EFL learners, the intrinsic load is relatively low for the present simple tense, as it is a grammatically

straightforward concept. In contrast, the intrinsic load increases for more complex structures, such as the passive voice or conditional sentences (Sweller, Ayres, & Kalyuga, 2011). The first type is intrinsic cognitive load, which is inherent to the content itself and cannot be reduced through design changes. The second type is extraneous cognitive load, which refers to how information is presented and the demands placed on the teaching process. When content is poorly designed, such as with unclear explanations, irrelevant images in short films, or expectations that learners must synthesize information from different sources, it creates a high level of extraneous cognitive load (Tarmizi & Sweller, 1988). This type of load wastes working memory capacity and does not contribute to effective learning.

The third variety of cognitive load, germane load, describes the constructive cognitive effort that learners put into processing information, constructing schema, and encoding this schema into LTM (Sweller et al., 2019). The goal of CLT-informed instructional design is to reduce extraneous load so WM resources can be allocated to germane processing. For example, when passive voice is covered using a short film, extraneous load can be minimized by bringing together a short explanation, a few clear examples, and a visual in a film exercise rather than spreading them around with clips or other materials or with other components. This strategy allows learners to focus cognitive load on learning the grammatical rule itself (germane load) rather than making their way through distracting or fragmented input (extraneous load).

Figure 4. *Cognitive Load Theory* (Paas et al., 2020)



Research corroborates the method: using audiovisual information through animated films, among others (Ammade & Ziska, 2021), contributes to improved clarity and reduction of unwarranted processing load. Using structured multimedia in civic and language lessons leads to improved outcomes achieved by learners as the instructional design is consistent with cognitive principles (Ardika, Agustiana, & Dibia, 2022) in both civic and language groups. This study of audiovisual media in adult education also emphasizes the reduction of extraneous stimuli for maximal learner engagement and schema formation (Nicolaou & Kalliris, 2020). Interactive pedagogical tools such as online quizzes also assist in delivering instruction without unnecessary stress and with direct focus on language skills (Oktaria, Rohmayadevi, & Murwantono, 2021). Furthermore, digital comic media can provide contextually focused input that reduces distractors and aids in germane processing (Susanti & Wibawa, 2022). Lastly, a technology-facilitated setting supported by high-level teacher self-efficacy has been shown to boost both student engagement and use of cognitive resources (Yang, 2024). Instructional grammar exercises are easier to structure and use in a strategy that is clear, direct, and within the wider cognitive architecture framework.

2.1.4 Depth Processing Hypothesis and Involvement Load Hypothesis

These modality-specific theories notwithstanding, general cognitive accounts of memory and retention support the efficacy of multimodal vocabulary instruction. Craik and Lockhart's (1972) Depth Processing Hypothesis states that retention is linked to the depth of initial processing used to process content; deeper semantic processing facilitates a stable trace of what is recalled compared to shallow perceptual processing. Multimodal tasks that require learners to link visual, auditory, and textual information stimulate deeper-level semantic thought. Similarly, Laufer and Hulstijn's (2001) Involvement Load Hypothesis posits that task-induced factors of need, search, and evaluation predict vocabulary retention. Multimodal tasks that require learners to find meaning through various methods and evaluate contextual appropriateness lead to a high involvement load that promotes retention.

2.2 Empirical Studies on Multimodal Instruction and Vocabulary Learning

The effectiveness of multimodal instruction for vocabulary learning has been extensively investigated across diverse EFL contexts, yielding a substantial body of empirical evidence that consistently supports the superiority of multimodal approaches over traditional text-only methods. This section reviews key studies that have examined various aspects of multimodal vocabulary instruction, including multimedia glossing, captioned video, annotated materials, and integrated multimedia packages.

2.2.1 Studies on Multimedia Glossing and Annotations

One of the most researched areas in multimodal vocabulary instruction is the use of multimedia glosses, annotations that provide definitions, translations, images, audio, or video to support word learning during reading or listening tasks. Chun and Plass (1996) conducted a foundational experiment investigating the effect of multimedia annotation on vocabulary learning in a computer-assisted language learning context. They found that learners who received visual and verbal annotations remembered significantly more words than those who received text-only glosses. The study demonstrated that the combination of text, pictures, and video annotations provided learners with multiple retrieval pathways, enhancing both immediate recognition and delayed recall.

Extending this line of inquiry, Plass et al. (1998) discovered that a multimedia instructional program with integrated visual and verbal explanations for unfamiliar words produced better performance on both immediate and delayed vocabulary tests compared to text-only instruction. Their findings emphasized that learners who could access both visual and verbal information were better able to construct coherent mental representations of new words, leading to more robust and durable vocabulary knowledge.

Yanguas (2009) investigated the contribution of different types of multimedia glosses, including textual glosses, pictorial glosses, and a mixture of both, on vocabulary learning and reading comprehension. The results indicated that the combined gloss condition produced the most vocabulary enhancement, consistent with the multimedia principle. Importantly, Yanguas found that pictorial glosses alone were more effective than textual glosses alone, but the combination of both yielded the highest gains, supporting the additive effect of dual coding.

More recently, Mahdi, Mohsen, and Almanea (2024) performed a second-round meta-analysis of multimedia gloss and vocabulary learning and observed a moderate overall effect size supporting multimedia over single-mode gloss, with greater benefit found when glosses included textual and pictorial input. The meta-analysis synthesized findings from numerous studies across various educational contexts, providing robust quantitative evidence for the effectiveness of multimedia glossing. Similarly, Ramezanali and Faez (2019) reported a meta-analytic effect between multimodal glossing and single-mode glossing in learning L2 vocabulary, which was strongest when a combination of visual and auditory modalities was used. Their analysis of 34 studies confirmed that multimedia glosses significantly outperform textual glosses alone, with effect sizes ranging from moderate to large depending on the modality combination.

In the Iraqi context specifically, Al-Seghayer (2001, 2016) conducted research that has been influential in shaping understanding of multimedia effects in vocabulary learning. While his studies were primarily situated in other contexts, his findings on the effectiveness of video annotations over still pictures or text-only annotations have direct relevance to Iraqi EFL settings. Al-Seghayer (2001) found that video clips, which provide dynamic contextual information and authentic language use, were significantly more effective than still images for vocabulary acquisition, as they conveyed temporal and motion information that static images could not capture.

2.2.2 Studies on Captioned and Subtitled Video

Beyond glossing, multiple multimodal teaching paradigms have been studied. According to Sydorenko (2010), students learned new vocabulary from both captioned and non-captioned videos, and combining audio, video, and text captions led to improvement in form recognition and meaning recall. Her research demonstrated that captioned video provides learners with simultaneous access to auditory input, visual context, and textual support, creating optimal conditions for vocabulary acquisition. Learners in the captioned video condition outperformed those who watched videos without captions on both form recognition and meaning recall measures.

Mohsen and Balakumar (2011) reviewed multimedia vocabulary learning studies and discovered that combining auditory and visual input is very beneficial in the vocabulary learning process. Their comprehensive review synthesized findings from over 50 empirical studies, concluding that the integration of audio and visual modalities consistently produced larger vocabulary gains than single-modality instruction. The review emphasized that the effectiveness of multimodal instruction depends on the quality and relevance of the materials, as well as the cognitive engagement of the learners.

Vanderplank (2016) conducted an extensive review of research on captioned and subtitled video in language learning, finding substantial evidence that captioned video improves vocabulary acquisition, listening comprehension, and overall language proficiency. His analysis highlighted that captions provide learners with a written representation of spoken language, facilitating form-meaning connections and supporting the development of both receptive and productive vocabulary knowledge. Vanderplank emphasized that the benefits of captioned video are particularly pronounced for intermediate learners, who can attend to both visual and auditory input while using captions as a support mechanism.

Further supporting these findings, Winke et al. (2010) investigated the effects of captioning on vocabulary learning and listening comprehension. Their study found that learners who viewed videos

with captions demonstrated significantly greater vocabulary gains than those who viewed videos without captions. The researchers also noted that the effectiveness of captioning varied depending on learner proficiency levels, with intermediate learners benefiting most from the combination of captions and audio.

2.2.3 Studies on Visual Aids, Pictures, and Illustrations

The role of visual aids in vocabulary learning has been extensively documented. Lin (2009) examined the effects of animated pictures on learning action verbs, finding that animation provided dynamic visual information that helped learners understand the meaning and usage of verbs more effectively than static images or text alone. The study demonstrated that animation reduced the cognitive effort required to infer meaning from context, allowing learners to allocate more cognitive resources to encoding and consolidation.

Al-Mansour and Al-Shorman (2012) investigated the effect of computer-assisted instruction on vocabulary learning among Saudi EFL learners, finding that learners who received instruction with visual aids significantly outperformed those who received traditional text-based instruction. Their study provided evidence that computer-assisted visual instruction facilitates vocabulary acquisition by presenting words in meaningful contexts with visual support.

Gass and Selinker (2008) discussed the role of visual input in second language acquisition, emphasizing that visual aids help learners create mental images that support vocabulary retention. They argued that pictures and illustrations provide learners with concrete referents for abstract words, facilitating the establishment of form-meaning connections. This is particularly important for learners in EFL contexts where exposure to authentic language is limited.

In a more recent study, Khoshshima et al. (2017) examined the effect of visual aids on vocabulary retention among Iranian EFL learners. Their findings revealed that learners who received vocabulary instruction with visual aids demonstrated significantly better retention than those who received text-only instruction. The study also found that the combination of visual aids with textual definitions produced the best results, supporting the dual coding principle.

2.2.4 Studies on Audio and Phonological Support

The auditory modality plays a crucial role in vocabulary learning, particularly in EFL contexts where learners have limited exposure to spoken English. Erfanian et al. (2016) investigated the effect of audio-supported vocabulary instruction on Iranian EFL learners' vocabulary retention. Their study found that learners who received audio input along with written definitions and example sentences significantly

outperformed those who received text-only instruction. The audio input provided learners with accurate pronunciation models and facilitated the development of phonological representations of new words, supporting both recognition and recall.

Rubin (1995) explored the role of audio in vocabulary acquisition, emphasizing that listening to words in context helps learners develop phonological awareness and facilitates form-meaning connections. Her research demonstrated that audio input, particularly when combined with visual support, provides learners with multiple opportunities to process and consolidate new lexical items.

In the Iraqi context, Hameed (2016) examined the effect of audio-visual materials on vocabulary learning among Iraqi EFL learners. The study found that learners who received instruction with audio-visual materials demonstrated significantly better vocabulary acquisition than those who received traditional text-based instruction. The researcher attributed these findings to the combined benefits of auditory and visual input, which provided learners with more comprehensive and meaningful exposure to new words.

2.2.5 Studies on Integrated Multimodal Packages

Recent research has increasingly focused on the effectiveness of integrated multimodal packages that combine multiple modalities, text, images, audio, and video, in a single instructional intervention. Rahimi and Allahyari (2016) investigated the effect of a multimedia vocabulary program on Iranian EFL learners' vocabulary knowledge and retention. The program incorporated written definitions, example sentences, audio pronunciations, images, and video clips for each target word. The results revealed that learners who used the multimedia program significantly outperformed those who received traditional vocabulary instruction on both immediate and delayed vocabulary tests.

Similarly, Khezrlou et al. (2017) examined the effects of computer-assisted language learning (CALL) on vocabulary acquisition and retention. Their study found that learners who received multimodal instruction through CALL programs demonstrated greater vocabulary gains and better retention than learners who received traditional instruction. The researchers attributed these findings to the multiple processing opportunities provided by multimodal input.

In a meta-analytic study, Zhang and Zou (2022) synthesized findings from 48 empirical studies on multimedia-assisted vocabulary learning, concluding that multimedia instruction generally produces large effects on vocabulary acquisition and moderate effects on vocabulary retention. Their analysis identified several factors that moderate the effectiveness of multimedia instruction, including the type of multimedia materials used, the proficiency level of the learners, and the duration of the intervention.

Notably, the meta-analysis found that multimedia instruction is particularly effective for intermediate learners, as they have sufficient foundational knowledge to benefit from the additional processing opportunities provided by multimodal input.

Lee and Mayer (2015) examined the effect of video-based multimedia instruction on vocabulary learning, comparing different presentation formats. Their study found that learners who received instruction with short video clips incorporating narration, visuals, and captions demonstrated significantly better vocabulary learning than those who received instruction with still images or text alone. The researchers emphasized the importance of temporal contiguity, presenting corresponding verbal and visual information simultaneously, in facilitating vocabulary learning.

Despite this abundant evidence from various EFL contexts, empirical research on multimodal vocabulary instruction remains notably sparse in the Iraqi EFL context. While studies conducted in other settings provide valuable insights, the unique contextual challenges faced by Iraqi learners, including limited authentic English exposure, resource constraints, large class sizes, and a predominantly examination-driven pedagogical culture, necessitate localized research to determine the effectiveness of multimodal approaches in this specific context. The present study addresses this gap by empirically investigating the effects of a comprehensive multimodal instructional package on vocabulary learning and retention among intermediate Iraqi EFL learners.

Although this is based on Western or East Asian EFL settings, this has further implications since Iraq also has very limited exposure to English and scarce resources, indicating an urgent need for instruction that maximizes the effectiveness of existing materials.

Though multimodal vocabulary instruction has ample theoretical and empirical support across the globe, it continues to receive very little scholarly attention in the Iraqi EFL context. The Iraqi system of learning is still largely based on a traditional teacher-oriented approach and, for the time being, focuses on translating grammar, series memorisation tasks, and language instruction mainly in Arabic (Saeed, 2015). Public schools lack resources like libraries, audio-visual tools, and additional materials to engage learners more effectively (Saeed, 2015), while the classroom experience is characterized by large class sizes, noise pollution, and a lack of infrastructure. These challenges are highlighted by the security situation being characterized as unstable, resulting in additional interruptions to the educational process and decreased student motivation to attend school (Saeed, 2015). This often leaves text-centered vocabulary learning ineffectual without providing learners with the rich contextualised language input needed for meaningful processing and long-term retention (Al-Marzouqi, 2020).

The present study is distinguished by its comprehensive approach to multimodal vocabulary instruction, simultaneously investigating the effects of four distinct modalities, written definitions, visual illustrations, audio recordings, and video clips, on both immediate vocabulary learning and delayed retention. While previous studies have typically examined individual modalities or limited combinations, this study provides a more holistic investigation of how integrated multimodal input affects vocabulary acquisition and retention in a resource-limited Iraqi EFL context. By drawing upon dual coding theory, cognitive theory of multimedia learning, and depth processing hypothesis, this research offers both theoretical validation and practical guidance for implementing multimodal vocabulary instruction in Iraqi classrooms.

Although research in other EFL settings has found that multimodal instruction can overcome a lack of L2 exposure and extend vocabulary acquisition by multiple sensory means (Chun & Plass, 1996; Yanguas, 2009; Mahdi et al., 2024), no empirical research has yet been conducted on the effect of a multimodal-oriented teaching programme on vocabulary learning and retention of Iraqi EFL learners to date. In view of the special social and contextual problems endured by Iraqi learners and the role of multimodal pedagogy in overcoming these difficulties, there is evident necessity for conducting a study in the field that addresses these points. Accordingly, the present study aimed to address the following research questions:

1. Does using multimodal instruction improve Iraqi EFL learners' immediate vocabulary learning compared to traditional instruction?
2. Does multimodal instruction significantly increase Iraqi EFL learners' vocabulary retention over two weeks compared to traditional instruction?

By addressing these questions, the study contributes to the limited literature by extending the understanding of technology-based and multisensory instruction within the Iraqi EFL context. These findings will help provide curriculum developers, teacher trainers, and classroom practitioners with practical guidance on effective multimodal vocabulary instruction that is sensitive to the resource constraints characteristic of Iraqi educational settings. Likewise, this investigation applies both well-established cognitive and multimedia learning theories to a relatively underexplored population, further enriching theoretical discussion on the modality effect in L2 vocabulary learning.

3. Methodology

3.1 Research Design

The study employs a quasi-experimental pre-test-post-test control group design to investigate the effects of multimodal instruction on vocabulary learning and retention. The independent variable is the type of vocabulary instruction, multimodal versus traditional, and the dependent variable is vocabulary learning measured by scores on an immediate post-test and vocabulary retention measured by scores on a delayed post-test administered two weeks after the intervention. This study consists of two groups: an experimental group receives vocabulary instruction through multiple modalities (written definitions, visual illustrations, audio recordings, and video clips), and a control group receives traditional text-based instruction. Both groups complete a pre-test, an immediate post-test after the six-week intervention, and a delayed post-test two weeks later.

The rationale for using multiple modalities in the experimental group is grounded in dual coding theory (Paivio, 1971) and cognitive theory of multimedia learning (Mayer, 2014), which predict that presenting vocabulary through multiple sensory channels creates redundant memory traces and facilitates deeper semantic processing, leading to better learning and retention. The use of four distinct modalities, written definitions, visual illustrations, audio recordings, and video clips, allows the study to capture the cumulative effect of multimodal input on vocabulary acquisition.

3.2 Participants

Sixty Iraqi EFL students enrolled in an intermediate-level English class at a private language school in Baghdad participated in the academic year 2024–2025. The learners were selected from an initial pool of 85 volunteers based on their performance on the Oxford Quick Placement Test (OQPT; Allan, 2004). Only participants scoring between 40 and 47, corresponding to level B1 (intermediate), were included in the study to ensure homogeneity. The final sample consisted of 32 male and 28 female participants aged 18–25 years. Participants were randomly assigned to the experimental group ($n = 30$) and the control group ($n = 30$). All individuals spoke Arabic as their first language and had at least seven years of English education in school. Data were collected from all participants after a comprehensive consent process.

3.3 Research Instruments

Three instruments were used in the research: the Oxford Quick Placement Test, a vocabulary test, and multimodal instructional materials.

The Oxford Quick Placement Test (OQPT; Allan, 2004) was used to assess general English proficiency and ensure group conformity among students. The test contains 60 items measuring grammar, vocabulary, and reading comprehension with a reported reliability of 0.90. The OQPT was selected because it is a well-established placement test with demonstrated validity and reliability in various EFL contexts.

The researcher developed a 30-item vocabulary test to assess knowledge of 30 target words from the intermediate-level syllabus used by the institute. The target content words, including nouns, verbs, and adjectives, were not previously taught to the participants. The test was administered in three distinct forms: Form A (pre-test), Form B (immediate post-test), and Form C (delayed post-test), with items arranged differently to minimize test-retest effects. Each item presented a target word in a sentence followed by four multiple-choice options: one clear synonym or short definition and three distractors. Content validity was confirmed by a panel of three experienced EFL teachers, and internal consistency was calculated using a pilot study with 15 learners of comparable proficiency, resulting in Cronbach's alpha of 0.87. The three parallel forms were designed to assess the same 30 target words but with different item orders and slightly varied sentence contexts, ensuring consistency in difficulty level while reducing practice effects.

Multimodal instructional materials were specifically created for the experimental group. Each target word was presented on a digital slide incorporating four elements: (a) a written definition in English, (b) a high-quality color photograph or illustration depicting the word's meaning, (c) an audio recording of a native speaker using the word in a sentence, and (d) a short video clip showing the word in context. These four modalities were intentionally selected to provide comprehensive sensory input that engages both verbal and non-verbal processing systems. Each learner received the materials on a tablet computer. Conversely, the control group received the same target words through paper materials, including only a written definition, an Arabic translation, and a sample sentence. This design allows for clear comparison between the multimodal package and traditional text-only instruction.

3.4 Data Collection Procedure

The study was conducted over eight weeks. In Week 1, volunteers were randomly assigned into groups, with 60 homogeneous intermediate learners identified after the OQPT was administered. The vocabulary pre-test (Form A) was given to both groups to assess prior vocabulary knowledge in Week 2. No significant difference in pre-test scores was found between the groups (see Section 4.1). The intervention began in Week 3 and was delivered over six weeks with sessions twice a week, each lasting

45 minutes (12 sessions total). Five target words were introduced in each session and then recycled in subsequent sessions.

The experimental group experienced multimodal instruction where each word was presented using a digital slide containing a written definition, a visual illustration, an audio recording of a native speaker using the word in a sentence, and a contextual video clip. This was followed by guided activities in which students engaged in small-group work, producing their own sentences linked to the word and image and discussing the video context. The multimodal presentation was designed so that learners could access all four modalities simultaneously, allowing them to process the information through both verbal and visual channels as predicted by dual coding theory (Paivio, 1971).

The control group received standard instruction: the teacher wrote each word on the whiteboard, introduced the Arabic translation and written definition, and required students to copy the information and complete written tasks, including gap-filling and translation drills. All 30 target words were covered by both groups during the sessions. To control for teacher effect, the same instructor taught all sessions.

In Week 9, one day following the last session, the immediate vocabulary post-test (Form B) was given to both groups. In Week 11, the delayed vocabulary post-test (Form C) was given to assess retention. Both post-tests were conducted in a controlled classroom without prior notice to the students to minimize preparation effects and capture natural learning outcomes.

The decision to use a two-week retention interval was guided by previous research on vocabulary retention (e.g., Yanguas, 2009; Ramezanali & Faez, 2019), which typically employs intervals ranging from one to four weeks to assess medium-term retention. The two-week interval was considered sufficient to distinguish between short-term memory effects and more durable vocabulary knowledge, while still being manageable within the study timeline. This interval also aligns with the practical needs of educational settings, where teachers need to know whether vocabulary gains are sustained beyond the immediate learning period.

3.5 Method of Data Analysis

Quantitative analysis was performed using SPSS software (version 27). Descriptive statistics (means and standard deviations) were calculated for pre-test, immediate post-test, and delayed post-test scores. The assumption of normality was evaluated using the Kolmogorov-Smirnov test. One-way analysis of covariance (ANCOVA) was conducted to compare group performance on the immediate and delayed post-tests while controlling for any pre-existing differences, with pre-test scores used as a covariate. Pre-

test scores were compared using an independent-samples t-test. Significance was evaluated at $\alpha = .05$. Effect sizes were computed based on partial eta squared (η^2) for ANCOVA and Cohen's d for t-tests.

The use of ANCOVA was considered appropriate for this study because it controls for pre-existing differences in vocabulary knowledge, increasing the statistical power and precision of the analysis. By using pre-test scores as a covariate, the ANCOVA adjusts for any initial differences between groups, providing a more accurate estimate of the treatment effect. The partial eta squared (η^2) effect sizes were interpreted according to Cohen's (1988) conventions: .01 = small, .06 = moderate, and .14 = large.

4. Results

4.1 Preliminary Analysis

Data were screened for normality and baseline equivalence before answering the research questions. The Kolmogorov-Smirnov test revealed that both pre-test, immediate post-test, and delayed post-test scores were normally distributed (all $p > .05$) for both groups. An independent samples t-test was performed to ensure that the difference between the experimental group ($M = 11.27$, $SD = 2.45$) and the control group ($M = 11.33$, $SD = 2.51$) was not statistically significant at the study start, $t(58) = 0.11$, $p = .915$. Therefore, prior vocabulary knowledge was relatively similar between the two groups.

4.2 Results for Research Question One

Research Question 1: Does multimodal instruction significantly improve immediate vocabulary learning for Iraqi EFL learners compared to traditional instruction? To address this question, a one-way ANCOVA was applied to immediate post-test scores with pre-test scores used as a covariate. Descriptive statistics for the immediate post-test are shown in Table 1.

Table 1. Descriptive Statistics for Immediate Vocabulary Post-Test

Group	N	Mean	SD	Adjusted Mean
Experimental	30	23.40	3.10	23.42
Control	30	18.73	3.45	18.71

The experimental group (Table 1) obtained a higher mean score on the immediate post-test (23.40) than the control group (18.73). The adjusted difference after controlling for pre-test scores remained significant. Table 2 summarizes the ANCOVA results.

Table 2. One-Way ANCOVA for Immediate Vocabulary Post-Test Scores

Source	SS	df	MS	F	p	η^2
Pre-test (Covariate)	88.12	1	88.12	10.45	.002	.152

Group	342.25	1	342.25	40.62	< .001	.413
Error	488.15	57	8.56			

The analysis demonstrated a statistically significant main effect of group, $F(1,57) = 40.62$, $p < .001$, with a very large effect size ($\eta^2 = .413$). The covariate (pre-test score) was also significant, $F(1,57) = 10.45$, $p = .002$. The adjusted mean of 23.42 for the experimental group and 18.71 for the control group indicates that after controlling for prior vocabulary knowledge, the multimodal instruction group achieved significantly better results than the traditional instruction group on the immediate vocabulary test. These findings suggest that multimodal instruction produced more intensive and immediate vocabulary learning.

4.3 Results for Research Question Two

Research Question 2: Does learning vocabulary via multimodal instruction increase vocabulary retention for Iraqi EFL students over two weeks compared to using standard instruction? One-way ANCOVA was also run on delayed post-test scores, taking pre-test scores as a covariate. Descriptive statistics are shown in Table 3.

Table 3. Descriptive Statistics for Delayed Vocabulary Post-Test

Group	N	Mean	SD	Adjusted Mean
Experimental	30	21.80	3.55	21.82
Control	30	14.87	3.90	14.85

The experimental group had a higher mean score on the delayed post-test (21.80) than the control group (14.87). ANCOVA results are shown in Table 4.

Table 4. One-Way ANCOVA for Delayed Vocabulary Post-Test Results

Source	SS	df	MS	F	p	η^2
Pre-test (Covariate)	102.41	1	102.41	8.96	.004	.134
Group	712.33	1	712.33	62.31	< .001	.521
Error	652.09	57	11.44			

A significant main effect of group was found, $F(1,57) = 62.31$, $p < .001$, with a very large effect size ($\eta^2 = .521$). The covariate was also significant, $F(1,57) = 8.96$, $p = .004$. The adjusted mean of 21.82 for the experimental group and 14.85 for the control group suggests that the multimodal group retained significantly more vocabulary than the control group after two weeks. These results indicate that multimodal instruction supported both initial learning and more durable longer-term retention.

Table 5. Score Changes from Immediate to Delayed Post-Test

Group	Immediate Mean	Delayed Mean	Mean Decrease	Percentage Retention
Experimental	23.40	21.80	1.60	93.2%
Control	18.73	14.87	3.86	79.4%

As shown in Table 5, the experimental group maintained 93.2% of their vocabulary knowledge over the two-week interval, compared to only 79.4% retention in the control group. This substantial difference in retention rates further demonstrates the durability of vocabulary learning achieved through multimodal instruction.

5. Discussion

This section presents our quantitative findings and their significance according to the theoretical framework and previous studies covered in the literature review. The research questions are restated, and results are discussed thematically, with interpretation informed by theoretical and empirical perspectives.

5.1 Discussion of Research Question One

The first research question addressed whether multimodal instruction enhances vocabulary learning in the short term compared to traditional teaching methods. Findings showed that the multimodal group significantly outperformed the traditional group on the immediate post-test with a large effect size ($\eta^2 = .413$). This result is in line with the prediction of Dual Coding Theory (Paivio, 1971), in which information processed through verbal and non-verbal channels forms richer memory representations. The multimodal materials allowed learners to engage with images, audio, and video in addition to written text, enabling them to encode new lexical items more frequently and helping to build a robust semantic network for each word.

Specifically, the digital slides used in this study provided learners with simultaneous access to four distinct types of information: (a) a written definition in English, which engaged the verbal processing system; (b) a high-quality color photograph or illustration, which engaged the visual processing system; (c) an audio recording of a native speaker using the word in a sentence, which provided phonological input and engaged the auditory processing system; and (d) a short video clip showing the word in context, which combined visual, auditory, and contextual information. This integration of multiple modalities created multiple memory traces for each vocabulary item. When learners attempted to recall a word later, they could access any of these traces, increasing the likelihood of successful retrieval. This

aligns with Paivio's (1986) assertion that dual coding, processing information through both verbal and non-verbal systems, produces an additive memory effect.

The effectiveness of each modality in the multimodal package can be understood through the lens of Mayer's (2014) multimedia principle, which suggests that individuals learn more effectively from words and pictures than from words alone. The visual illustrations and video clips in the experimental condition provided concrete referents for abstract concepts, making the meanings of new words more accessible and memorable. The audio recordings provided accurate pronunciation models and helped learners establish phonological representations of new words, which is particularly important in EFL contexts where learners have limited exposure to spoken English (Erfanian Mohammadi et al., 2016).

The video clips, which showed the target words in authentic or semi-authentic contexts, were especially valuable because they provided rich contextual information that helped learners understand not only the meaning but also the appropriate usage of each word. Research by Al-Seghayer (2001) has shown that video annotations are more effective than still images for vocabulary learning because they convey temporal and motion information that static images cannot capture. The video clips in this study also provided examples of how the words are used in natural discourse, supporting the development of collocational and grammatical knowledge.

The active processing assumption of Mayer's (2014) theory was also upheld, as learners chose appropriate visual and auditory information from the slides, organized it into coherent mental structures, and integrated it with their existing knowledge during small-group activities. The experimental group's activities, discussing video contexts, producing their own sentences, and connecting words to images, required deep engagement with the material, consistent with the generative processing principle of CTML. These tasks stimulated deeper semantic processing, which Craik and Lockhart's (1972) Depth Processing Hypothesis identifies as a vital component for establishing long-term memory.

The involvement load created by the multimodal tasks also contributed to the strong learning outcomes. According to Laufer and Hulstijn's (2001) Involvement Load Hypothesis, tasks that require learners to search for meaning and evaluate contextual appropriateness lead to higher involvement load and better retention. In the experimental group, learners were asked to find meaning in images, compare the word across different modalities, and evaluate the appropriateness of word meanings in various contexts. These processes required greater semantic processing and created more elaborate memory traces, supporting both immediate learning and longer-term retention.

Findings also concur with Yanguas (2009), who observed the highest vocabulary gains using a combination of textual and pictorial glosses. The present study builds on this outcome using a wider multimodal design with audio and video alongside static imagery. The large effect size found in this study ($\eta^2 = .413$) indicates that the benefits of multimodal instruction are substantial even in a resource-limited EFL context like Iraq. This finding is consistent with meta-analytic findings by Mahdi et al. (2024), who reported moderate to large effects for multimedia glosses on vocabulary learning, and Ramezanali and Faez (2019), who found that multimodal glossing produced stronger effects than single-mode glossing.

The pedagogical implications of these findings are significant. In the Iraqi EFL context, where learners have limited exposure to English outside the classroom and resources are constrained, multimodal instruction offers a viable and effective approach to vocabulary teaching. The use of digital slides incorporating multiple modalities, even in resource-limited settings, can be implemented using basic technology such as tablets, smartphones, or even projectors. The key is to ensure that the multimodal materials are well-designed and aligned with cognitive principles of learning.

5.2 Discussion of Research Question Two

The second research question addressed whether multimodal instruction increases vocabulary retention over time. The multimodal group performed significantly better than the control group on the delayed post-test, with a very large effect size ($\eta^2 = .521$). The experimental group had a mean decrease of only 1.60 points from the immediate to delayed post-test, while the control group had a mean decrease of 3.86 points, indicating that multimodal instruction supported more sustained retention.

These retention benefits can be attributed to several interconnected theoretical mechanisms. First, the dual coding effect (Paivio, 1971) provided learners with multiple retrieval options. When attempting to recall a word, learners could access either a verbal or a visual representation, or both, enhancing the likelihood of successful recall despite the delay. The presence of multiple memory traces, one for each modality, created redundancy that protected against forgetting. This is consistent with Paivio's (1986) assertion that dual coding creates "redundant memory traces" that are more resistant to decay over time.

Second, the multimodal activities promoted deeper processing (Craik & Lockhart, 1972), resulting in more elaborate and distinctive memory traces that are less prone to forgetting. The experimental group's engagement with multiple modalities, reading definitions, viewing illustrations, listening to audio, watching video clips, and discussing contexts, required deeper semantic processing than the control group's rote memorisation and translation activities. According to Craik and Lockhart (1972), deeper

processing, which involves semantic and elaborative encoding, produces more durable memory traces than shallow perceptual processing.

Third, the reduction of extraneous cognitive load and facilitation of germane processing, as suggested by Cognitive Load Theory (Sweller, 1988), enabled learners to spend a greater share of their cognitive resources constructing and preserving schemas, generating more organised and retrievable lexical knowledge. The multimodal materials were designed to reduce extraneous load by presenting information in a clear, integrated format that minimized cognitive distractions. This allowed learners to focus their cognitive resources on germane processing, constructing meaningful associations between new words and existing knowledge structures.

Fourth, the multimodal instruction provided learners with rich contextual information that supported the development of lexical networks. The video clips and illustrations provided situational and visual contexts for each word, helping learners understand how the word is used in authentic situations. This contextual information facilitated the formation of semantic connections between new words and existing knowledge, making the lexical knowledge more integrated and accessible.

Retention results are consistent with previous research. In their meta-analysis, Ramezanali and Faez (2019) found that multimodal glosses produced superior performance on delayed post-tests, providing evidence that multimodal encoding improves long-term retention. Sydorenko (2010) also demonstrated that vocabulary gains from captioned video were sustained over time. Building on this evidence, the current research shows that a multimodal package with multiple media elements, images, audio, video, and text, can provide more sustained vocabulary storage in an Iraqi EFL setting where learners lack adequate exposure to English outside the classroom and depend on instruction for lexical development.

The retention benefits of multimodal instruction are particularly significant in the Iraqi EFL context. Traditional instruction, which emphasizes rote memorisation and translation, tends to produce vocabulary knowledge that is fragile and easily forgotten. In contrast, multimodal instruction creates robust lexical representations that are more resistant to forgetting, as evidenced by the experimental group's superior retention of 93.2% of vocabulary knowledge over two weeks compared to 79.4% retention in the control group. This finding has important implications for curriculum design and teaching practice, suggesting that investment in multimodal materials may yield long-term benefits that justify the initial resource costs.

The study also contributes to theoretical understanding of vocabulary acquisition in EFL contexts. The results confirm that the principles of dual coding theory and cognitive theory of multimedia learning

apply not only to well-resourced educational settings but also to resource-constrained environments like Iraqi classrooms. The findings extend the generalizability of these theoretical frameworks to a new population and context, enriching the theoretical discussion on modality effects in L2 vocabulary learning.

6. Conclusion

This research aimed to investigate the significance of multimodal instruction in learning and retaining vocabulary among intermediate Iraqi EFL learners. The findings confirmed that learners using vocabulary instruction integrating visual, auditory, and written sources were significantly superior after instruction and two weeks post-instruction compared to traditional text-based instruction. The large effect sizes on both learning and retention demonstrate both practical and theoretical significance of the multimodal model.

6.1 Summary of Findings

The study produced three main findings:

First, the multimodal group ($n = 30$) significantly outperformed the control group ($n = 30$) on the immediate vocabulary post-test ($F(1,57) = 40.62, p < .001, \eta^2 = .413$). This demonstrates that the integrated use of written definitions, visual illustrations, audio recordings, and video clips produced significantly better immediate vocabulary learning than traditional text-based instruction.

Second, the multimodal group demonstrated significantly better vocabulary retention over a two-week period ($F(1,57) = 62.31, p < .001, \eta^2 = .521$). The experimental group maintained 93.2% of their vocabulary knowledge, compared to only 79.4% retention in the control group, indicating that multimodal instruction produces more durable lexical knowledge.

Third, the large effect sizes for both immediate learning and retention indicate that multimodal instruction has substantial practical significance in the Iraqi EFL context. These effects were consistent across both research questions, suggesting that the benefits of multimodal instruction are robust and reliable.

6.2 Pedagogical Implications

The findings of this study have several important pedagogical implications for Iraqi EFL professionals and educators in similar resource-constrained contexts:

Table 6. Pedagogical Implications for Multimodal Vocabulary Instruction in Iraqi EFL Classrooms

Modality	Instructional Application	Theoretical Basis	Implementation Strategy
Written Definitions	Provide clear, comprehensible definitions in English	Verbal processing (DCT)	Use simple language, include example sentences, avoid overly complex definitions
Visual Illustrations	Use photographs, drawings, or diagrams to represent word meanings	Visual processing (DCT)	Ensure images are high-quality, relevant, and culturally appropriate
Audio Recordings	Provide native speaker pronunciations and usage examples	Auditory processing (DCT, CTML)	Include multiple examples of words in varied contexts
Video Clips	Show words in authentic or semi-authentic contexts	Dynamic contextual learning (CTML)	Use short clips (30-60 seconds), ensure clear demonstration of word meaning and usage

For Iraqi EFL professionals, the incorporation of multimodal resources, even simple digital slides on multimedia mediums (e.g., images, audio, and brief videos), can greatly enhance vocabulary instruction. The specific findings regarding each modality suggest that:

The written definitions provide a foundation for verbal processing, but should be supplemented with visual and auditory input to create dual coding effects. The visual illustrations help learners form mental images that support memory and understanding, particularly for concrete nouns and actions. The audio recordings provide accurate pronunciation models, which are crucial for developing receptive and productive vocabulary knowledge. The video clips offer rich contextual information that helps learners understand how words are used in authentic situations, supporting the development of collocational and grammatical knowledge.

For curriculum designers and material developers, the findings suggest that integrating “multimodal” elements into textbooks and resources, especially in schools with limited exposure to real English, can significantly improve vocabulary learning and retention. This may involve creating supplementary materials that provide visual, auditory, and contextual support for vocabulary items. In settings with limited technology, even simple materials such as printed images or audio recordings played on a speaker can provide multimodal input.

For teacher training, the findings highlight the need to prepare teachers to design multimodal activities that are compatible with multimedia learning standards. Teachers should be trained to use

multimedia materials in ways that reduce rather than detract from cognitive load. This includes following Mayer's (2014) principles for effective multimedia design, such as eliminating irrelevant information, presenting corresponding verbal and visual information simultaneously, and segmenting complex information into manageable steps.

For classroom practitioners, the findings offer practical guidance on implementing multimodal vocabulary instruction. Teachers can incorporate multimodal activities into their lessons by: (a) using images and real objects to illustrate word meanings; (b) playing audio recordings of words and sentences; (c) showing short video clips that demonstrate word usage; and (d) designing tasks that engage multiple modalities simultaneously, such as watching a video with captions and writing sentences using the target vocabulary.

Furthermore, given the existing resource constraints identified in Iraqi public schools (Saeed, 2015), low-cost solutions such as using mobile devices to display multimodal content may be effective for instructional improvement. Smartphones, tablets, and even projectors can be used to present multimodal materials at minimal cost. Teachers can also create their own multimodal materials using free or low-cost digital tools, such as creating audio recordings with a smartphone or finding relevant images and videos online.

The practical significance of these findings is substantial. In a context where resources are limited and traditional instruction often fails to produce lasting vocabulary knowledge, multimodal instruction offers a viable and effective alternative. The large effect sizes observed in this study indicate that even modest investments in multimodal materials can produce significant improvements in vocabulary learning and retention.

6.3 Contributions to Knowledge

This study makes several important contributions to knowledge in the field of L2 vocabulary acquisition:

First, it provides empirical evidence for the effectiveness of multimodal vocabulary instruction in the Iraqi EFL context, filling a significant gap in the literature. While multimodal approaches have been extensively studied in other contexts, this is one of the first studies to systematically investigate their effects in Iraq.

Second, it extends the application of dual coding theory (Paivio, 1971), cognitive theory of multimedia learning (Mayer, 2014), and depth processing hypothesis (Craik & Lockhart, 1972) to a new

population and context. The findings confirm that these theoretical frameworks are applicable not only to well-resourced educational settings but also to resource-constrained environments.

Third, it provides empirical evidence for the specific contributions of different modalities, written definitions, visual illustrations, audio recordings, and video clips, to vocabulary learning and retention. This detailed analysis of modality effects offers insights that can inform the design of multimodal materials and instruction.

Fourth, it offers practical guidance for implementing multimodal vocabulary instruction in resource-constrained contexts. The study demonstrates that effective multimodal instruction does not require expensive technology or sophisticated equipment; even basic multimodal materials can produce significant learning gains.

6.4 Limitations of the Study

The study also has several limitations that should be acknowledged. First, the sample was drawn from one private language school with only intermediate-level learners, which may impact the generalizability of the findings to novice, advanced, or students in public school contexts. Future research should include learners at different proficiency levels and from various institutional settings.

Second, the intervention was relatively brief, comprising only six weeks, and the retention interval was only two weeks, so the findings do not reflect longer-term retention. Longitudinal studies with extended retention periods would provide a more complete understanding of the durability of multimodal learning effects.

Third, the multimodal materials used a particular combination of image, audio, and video; different combinations of modalities might produce different results. Future research should systematically compare different modality combinations to determine the most effective configurations for vocabulary teaching.

Fourth, the study used only receptive multiple-choice vocabulary tests, which do not measure learners' productive use of target words. Future research should include productive measures, such as writing tasks or oral production, to assess whether multimodal instruction supports productive vocabulary knowledge.

Fifth, the study did not examine individual difference variables such as learning style, working memory capacity, or motivation. Future research should investigate how these variables moderate the effects of multimodal instruction, as different learners may benefit differently from various modality combinations.

6.5 Recommendations for Future Research

Based on the findings and limitations of this study, several recommendations for future research are offered:

First, the study should be replicated with larger and more diverse samples at different proficiency levels and institutional contexts in Iraq. This would enhance the generalizability of the findings and allow for sub-group analyses to identify which learners benefit most from multimodal instruction.

Second, longitudinal studies with extended retention periods (e.g., 3-6 months) would provide a more complete understanding of the sustainability of multimodal learning effects. This is important for determining whether the retention benefits observed in this study persist over longer periods.

Third, research should systematically compare different combinations of modalities to identify the most effective configurations for vocabulary teaching. For example, video versus still images, audio versus text, and different combinations of these modalities could be compared to determine the optimal multimodal package.

Fourth, research should include productive measures of vocabulary knowledge, such as writing tasks or oral production, to assess the full range of multimodal instruction effects. This would provide a more comprehensive understanding of how multimodal instruction supports vocabulary development.

Fifth, future research should investigate the effects of individual difference variables, such as learning style, working memory capacity, and motivation, on the effectiveness of multimodal instruction. Understanding how different learners respond to multimodal instruction could inform the development of differentiated instructional approaches.

Sixth, research should explore the integration of multimodal instruction with other pedagogical approaches, such as task-based language teaching, to examine whether multimodal elements can enhance the effectiveness of these approaches. This would contribute to the development of more comprehensive and effective instructional models.

Finally, future research should investigate the cost-effectiveness of multimodal instruction in resource-constrained settings, examining the relationship between investment in multimodal materials and learning outcomes. This would provide practical guidance for educational decision-makers on how to allocate limited resources effectively.

6.6 Concluding Remarks

In conclusion, this study provides strong empirical evidence for the effectiveness of multimodal vocabulary instruction in the Iraqi EFL context. The findings confirm that integrating written

definitions, visual illustrations, audio recordings, and video clips produces significantly better vocabulary learning and retention compared to traditional text-based instruction. The large effect sizes observed indicate that multimodal instruction has substantial practical significance, offering a viable and effective approach to vocabulary teaching in resource-constrained settings.

The theoretical contributions of the study are also significant. The findings provide empirical support for dual coding theory, cognitive theory of multimedia learning, and depth processing hypothesis in a new population and context, enriching theoretical discussion on modality effects in L2 vocabulary learning. The study demonstrates that these theoretical principles are applicable across diverse educational settings, supporting the universality of cognitive learning principles.

The pedagogical implications are clear: Iraqi EFL teachers, curriculum designers, and material developers should embrace multimodal approaches to vocabulary instruction. Even with limited resources, simple multimodal materials can significantly improve vocabulary learning and retention. The key is to ensure that multimodal materials are well-designed and aligned with cognitive principles of learning.

Ultimately, this study contributes to improving EFL vocabulary instruction in Iraq by providing evidence-based guidance for implementing effective multimodal approaches. By addressing the challenges of limited authentic English exposure and resource constraints, multimodal instruction offers a pathway to more effective vocabulary teaching and learning in Iraqi educational settings.

References

- Allan, D. (2004). *Oxford Placement Test 1: Test Pack*. Oxford University Press.
- Al-Mansour, N. S., & Al-Shorman, R. A. (2012). The effect of computer-assisted instruction on Saudi university students' learning of English. *Journal of King Saud University - Languages and Translation*, 24(1), 51-57.
- Al-Marzouqi, A. H. (2020). The impact of rote learning and translation methods on vocabulary acquisition among Iraqi EFL learners. *Journal of University of Babylon for Humanities*, 28(3), 45–57.
- Al-Seghayer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning & Technology*, 5(1), 202–232.
- Al-Seghayer, K. (2016). Celebrating the story of my first contribution to CALL. *Language Learning & Technology*, 20(3), 179–184.

- Ammade, S., & Ziska, I. (2021). Animation film in writing teaching instruction on narrative text: Working or not working? *Language Circle: Journal of Language and Literature*, 16(1), 81–88.
- Ardika, N., Agustiana, I., & Dibia, I. (2022). Character and learning outcomes of civics in VCT learning assisted by audio visual media. *Indonesian Values and Character Education Journal*, 5(1), 19–31.
- Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology*, 98(1), 182–197.
- Butcher, K. R. (2014). The multimedia principle. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 174–205). Cambridge University Press.
- Chun, D. M., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. *The Modern Language Journal*, 80(2), 183–198.
- Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 3(3), 149–210.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 87–114.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671–684.
- Erfanian Mohammadi, J., Khalaji, H., & Naderi, F. (2016). The effect of audio-supported vocabulary instruction on Iranian EFL learners' vocabulary retention. *International Journal of English Linguistics*, 6(3), 178-188.
- Gass, S. M., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd ed.). Routledge.
- Hameed, A. F. (2016). The effect of audio-visual materials on vocabulary learning among Iraqi EFL learners. *Journal of Educational Sciences*, 27(4), 45-62.
- Herrington, J., & Oliver, R. (2000). An instructional design framework for authentic learning environments. *Educational Technology Research and Development*, 48(3), 23–48.
- Kanellopoulou, C., Kermanidis, K., & Giannakouloupoulos, A. (2019). The dual coding theory and its applications in language learning. *International Journal of Language and Linguistics*, 7(3), 124-135.
- Khoshsima, H., Rezaeian, H., & Rezaeian, M. (2017). The effect of visual aids on vocabulary retention among Iranian EFL learners. *Journal of Language Teaching and Research*, 8(4), 765-772.

- Khezrlou, S., Ellis, R., & Sadeghi, K. (2017). Effects of computer-assisted language learning on vocabulary acquisition and retention. *Computer Assisted Language Learning*, 30(3-4), 283-304.
- Koscianski, A., & Zanutto, D. D. C. F. (2014). A design model for educational multimedia software. *Creative Education*, 5(22), 2003–2016.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1–26.
- Lee, H., & Mayer, R. E. (2015). Video-based multimedia instruction for vocabulary learning. *Educational Technology Research and Development*, 63(5), 715-738.
- Lin, C.-C. (2009). Learning action verbs with animation. *JALT CALL Journal*, 5(1), 23–40.
- Mahdi, H. S., Mohsen, M. A., & Almanea, M. (2024). Multimedia glosses and second language vocabulary learning: A second-round meta-analysis. *Acta Psychologica*, 248, Article 104341.
- Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist*, 32(1), 1–19.
- Mayer, R. E. (2001). *Multimedia learning*. Cambridge University Press.
- Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). Cambridge University Press.
- Mayer, R. E. (2008). Applying the science of learning: Evidence-based principles for the design of multimedia instruction. *American Psychologist*, 63(8), 760–769.
- Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 43–71). Cambridge University Press.
- Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 279–315). Cambridge University Press.
- Moeser, S. D., & Bregman, A. S. (1973). Imagery and language acquisition. *Journal of Verbal Learning and Verbal Behavior*, 12(1), 91–98.
- Mohsen, M. A., & Balakumar, M. (2011). A review of multimedia glosses and their effects on L2 vocabulary acquisition. *English Language Teaching*, 4(1), 199–208.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Nation, I. S. P. (2006). *How vocabulary is learned*. Oxford University Press.

- Nicolaou, C., & Kalliris, G. (2020). Audiovisual media communications in adult education: The case of Cyprus and Greece of adults as adult learners. *European Journal of Investigation in Health Psychology and Education*, 10(4), 967–994.
- Oktaria, A., Rohmayadevi, L., & Murwantono, D. (2021). Online game quiz “Kahoot” in teaching English for students of SMP Muhammadiyah Yogyakarta. *Project: Professional Journal of English Education*, 4(2), 290–296.
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38(1), 1–4.
- Paas, F., Kalyuga, S., & Renkl, A. (2020). *Cognitive load theory: Instructional implications for instructional design*. Routledge.
- Paivio, A. (1971). *Imagery and verbal processes*. Holt, Rinehart & Winston.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, 45(3), 255–287.
- Paivio, A. (2014). Intelligence. In *Elsevier encyclopedia of intelligence* (Vol. 47, pp. 141–158). Elsevier.
- Paivio, A., & Begg, I. (1981). *Psychology of language*. Prentice-Hall.
- Paivio, A., & Desrochers, A. (1980). A dual-coding approach to bilingual memory. *Canadian Journal of Psychology*, 34(4), 388–399.
- Paivio, A., Walsh, M., & Bons, T. (1994). Concreteness effects on memory: When and why? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(5), 1196–1204.
- Park, I., & Hannafin, M. (1993). Empirically-based guidelines for the design of interactive multimedia. *Educational Technology Research and Development*, 41(1), 63–85.
- Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second-language multimedia learning environment. *Journal of Educational Psychology*, 90(1), 25–36.
- Pylyshyn, Z. W. (1973). What the mind’s eye tells the mind’s brain: A critique of mental imagery. *Psychological Bulletin*, 80(1), 1–25.
- Rahimi, M., & Allahyari, H. (2016). The effect of a multimedia vocabulary program on Iranian EFL learners’ vocabulary knowledge and retention. *Journal of Language Teaching and Research*, 7(6), 1213–1221.

- Ramezanali, M., & Faez, F. (2019). Efficacy of multimodal glossing on L2 vocabulary learning: A meta-analysis. *TESOL Quarterly*, 53(4), 1026–1057.
- Reed, S. K. (2010). *Cognition: Theories and applications* (8th ed.). Wadsworth Cengage Learning.
- Rieber, L. P. (1994). *Computers, graphics, and learning*. Brown and Benchmark.
- Rubin, J. (1995). The contribution of audio to vocabulary acquisition. *Language Learning*, 45(4), 681-704.
- Rusanganwa, J. A. (2015). Developing a multimedia instrument for technical vocabulary learning. *Computer Assisted Language Learning*, 28(1), 97–111.
- Saeed, N. R. (2015). *Problems of teaching English in Iraqi public schools in Mosul for the intermediate stage from the teachers' point of view* [Unpublished master's thesis]. Middle East University, Amman, Jordan.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan.
- Soh, K. C. (2010). Bilingual dual-coding and code-switching: Implications for the L1 in L2 learning. *Journal of Linguistics and Language Teaching*, 1(2), 271–296.
- Sternberg, R. J. (2003). *Cognitive psychology* (3rd ed.). Thomson/Wadsworth.
- Susanti, K., & Wibawa, I. (2022). Learning socio-cultural diversity of society containing tolerance character values through digital comic media. *Mimbar PGSD Undiksha*, 10(2), 342–350.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285.
- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19–30). Cambridge University Press.
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer.
- Sweller, J., Ayres, P., & Kalyuga, S. (2019). *Cognitive load theory* (2nd ed.). Springer.
- Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning & Technology*, 14(2), 50–73.
- Tarmizi, R. A., & Sweller, J. (1988). Guidance during problem solving. *Journal of Educational Psychology*, 80(4), 424–434.
- Vanderplank, R. (2016). Captioned and subtitled video in language learning. *The Modern Language Journal*, 100(S1), 204-219.
- Wilkins, D. A. (1972). *Linguistics in language teaching*. Edward Arnold.
- Winke, P., Gass, S. M., & Sydorenko, T. (2010). The effects of captioning on vocabulary learning and listening comprehension. *Language Learning & Technology*, 14(3), 65-89.

- Wittrock, M. C. (1974). Learning as a generative process. *Educational Psychologist*, 11(2), 87–95.
- Yang, Z. (2024). Examining the relationship between teacher self-efficacy and student engagement in technology-enhanced learning environments. *International Journal of New Developments in Education*, 6(2).
- Yanguas, I. (2009). Multimedia glosses and their effect on L2 text comprehension and vocabulary learning. *Language Learning & Technology*, 13(2), 48–67.
- Zhang, D., & Zou, D. (2022). Multimedia-assisted vocabulary learning: A meta-analysis. *Computer Assisted Language Learning*, 35(8), 1894-1924.

Author Information

**Rusul Esmaeel Sakran (General Directorate of
Education in Wasit Governorate , Ministry of
Education)**
goldenrose84627@gmail.com
