An-Najah National University
Faculty of Graduate Studies

THE EFFECT OF AI ON ENGLISH LANGUAGE RESEARCH QUALITY OF PALESTINIAN GRADUATE STUDENTS

By
Amani Adnan Naeem Abed

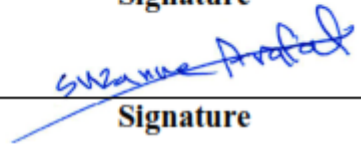
Supervisor
Dr. Nedal Jayousi

This Thesis is Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Methods of Teaching English, Faculty of Graduate Studies, An-Najah National University, Nablus, Palestine.
2025

THE EFFECT OF AI ON ENGLISH LANGUAGE RESEARCH QUALITY OF PALESTINIAN GRADUATE STUDENTS

By
Amani Adnan Naeem Abed

This Thesis was Defended Successfully on 17/12/2025 and approved by:

Dr. Nedal Rashed Jayousi **Supervisor**	 **Signature**
Dr. Mohamad Salah Harzallah **External Examiner**	 **Signature**
Dr. Suzanne Arafat **Internal Examiner**	 **Signature**

Dedication

First…

All praise is due to Allah, always and forever. He is the One who, when He wills something, simply says "Be," and it is. To Him belongs the grace of placing love for knowledge in my heart and easing my steps along this path. My Lord, thank You for preparing the reasons and circumstances that allowed me to walk the road I always prayed would be easy. The journey was difficult… but with patience, Your mercy, and Your kindness, every hardship became lighter, and by Your grace, I was saved.

To the first embrace… the warmest and most faithful…To my first teacher and role model, my dear father, Mr. Adnan Abed, Thank you for setting me on the right path, for every bit of support you have given me, and for every feeling of safety and strength you planted within me. Words fall short of expressing what you mean to me.

And to my beloved mother, Huda Abed, Thank you for your support and for the prayers that have always protected me. To my dear parents, the word "thank you" is not enough; this achievement is from you and returns to you.

To my dear husband, Farid Jaradat …Thank you for being my supporter, my strength, and the space of love and tenderness in my life. Everything becomes easier with you by my side.

To my beloved children…

To Heaven—my sweet eldest daughter, beautiful inside and out; patient, aware, and wise beyond her years; my companion in the most important stages of my life and the owner of the kindest and most tender heart—thank you, my love. To Adam and Liana, the most beautiful twins, thank you for making everything easier just by being here. You are part of my success and my strength in every step.

To my brothers and sisters…I dedicate this accomplishment to you in gratitude for your guidance, your support, and your constant encouragement.

Acknowledgment

First and foremost, all praise and gratitude are due to Allah, whose grace, guidance, and mercy have accompanied me throughout this journey. This work is not the result of individual effort; rather, it represents the collective support, collaboration, and contribution of many individuals and institutions who played an essential role in bringing this study to completion.

I would like to express my deepest appreciation to my supervisor, Dr. Nidal Al-Jayousi, whose guidance, sincere dedication, and valuable insights had a profound impact on my master's journey. His continuous encouragement and academic support were instrumental in shaping this research.

My sincere thanks also go to the administration and academic staff of An-Najah National University for providing a supportive academic environment and for their consistent assistance throughout all stages of my study.

I am equally grateful to the Student Affairs Department for granting me the official approval to distribute the questionnaire across universities in the West Bank, which enabled me to reach a diverse and representative sample crucial for the success of this research.

Special thanks are extended to the universities that kindly facilitated the distribution of the questionnaire, including An-Najah National University, Bethlehem University, Birzeit University, Al-Quds Open University, Al-Quds University – Abu Dis, the Arab American University, Palestine Technical University–Kadoorie, the University of Hebron, and Palestine Polytechnic University. Their cooperation and collective effort greatly supported the completion of this study.

It also goes without reflection that my deepest thanks and gratitude must be given to the Master's Thesis Committee: the external examiner, Dr. Mohamad salah Herzallah, and the internal examiner, Dr. Suzan Arafat, whose time, valuable comments, and constructive feedback contributed significantly to enhancing the quality and clarity of this research.

To everyone who contributed, supported, or encouraged me during this academic journey—thank you. Your efforts and cooperation have truly made a meaningful difference.

Declaration

I, the undersigned, declare that I submitted the thesis entitled:

THE EFFECT OF AI ON ENGLISH LANGUAGE RESEARCH QUALITY OF PALESTINIAN GRADUATE STUDENTS

I, declare that the work provided in this thesis, unless otherwise referenced, is the researcher's own work, and has not been submitted elsewhere for any other degree or qualification.

Student's Name:	**Amani Adnan Naeem Abed**
Signature:	*Amani abed*
Date:	**17/12/2025**

List of Contents

List of Tables

List of Appendices

THE EFFECT OF AI ON ENGLISH LANGUAGE RESEARCH QUALITY OF PALESTINIAN GRADUATE STUDENTS

By
Amani Adnan Naeem Abed
Supervisor
Dr. Nedal Jayousi

Abstract

This study aimed to investigate the effect of artificial intelligence (AI) tools on the quality of English-language academic research as perceived by Palestinian graduate students. Due to the lack of local research on the impact of AI on academic work, this study revealed students' perspectives on using AI tools in research writing, focusing on their strengths, weaknesses, and strategies for maximize use.

A descriptive quantitative approach was selected, and data were collected through a questionnaire distributed electronically. The study sample included 171 valid responses from Palestinian graduate students.The instrument assessed high reliability, with a Cronbach's Alpha coefficient of 0.907. Data analysis indicated that students generally had positive perceptions regarding the effect of AI tools in academic contexts. The overall mean score across the three domains reached 3.58 (71.6%), indicating a high level of perceived effectiveness. Participants particularly appreciated AI tools for saving time, enhancing language accuracy, and organizing information, while issues were raised regarding credibility, limited access, and AI's difficulty in understanding context.

The study recommended adapting a balanced approach that combines AI support with students' own efforts. It also referred the significant of providing training sessions on ethical and effective AI use and increasing institutional reach to AI tools within higher education.

Keywords: Artificial Intelligence, Research Quality, Graduate Students, Academic Writing, Palestine.

Chapter One

Introduction and Theoretical Background

1.1 Introduction

Recently artificial intelligence (AI) has begun to have profound impact on the education sector in general and academic research in particular.Adapting AI tools In many universities become urgent necessary to assist students in duties such as in analyzing information and arranging their research. Its importance lies in providing quick feedback, language assessment and also support in modifying texts. This technology is prevalent in Western universities, while Palestinian universities still in need for careful attention and care regarding using AI tools.

Palestinian universities have widened their programs to create new opportunities for graduate student despite ongoing political and economic challenges. Related English programmers,in particular, gained special attention from universities since it is dominated language for global research and scholarly publishing. However, students encounter challenges when they are required to produce written work in English. Students and colleges have determined that graduate students' writing problems represent the main obstacles that reduce the quality of research writing in English. in the west bank.

Absence proficiency of academic writing standers is considered one of the most obvious obstacle, especially when time comes to organize debates, paraphrase sentences, and keep coherence. Also considering literary translation from Arabic to English as another limitation since it leads to lack of clarity in students' work. Moreover, in Palestinian universities accessing the academic writing support center or providing workshop training is limited.All of previous restrictions rise tension because students are worry about improving their skills independently. This situation is more problematic because of the lack of updated resources and accessing internet that sometimes prohibit students from Gaining access to global databases and journals (Hamdan, Ashour, & Daher, 2021).

The situation has become increasingly strained due to supervisors insistence on students to conduct research of a quality that aligns with international standards.This enhance the gap between expected performance and students' actual capacity within limited resources. As (Smith & Scott, 2023) & stated that is lead to anxiety among students, causing them

to postpone completing thesis and repeatedly revise their work multiple times before it is accepted. Recent studies also show that students in Palestinian universities suffer with combined critical thinking into their writing because most of their previous training centered on Sentence structure and vocabulary rather than Scholarly discussion and debate (Badah, Najjar, Almimi, & Alqbailat, 2024), these challenges suggest that English academic writing is hard not only because language itself, it's also about its structure and rules. If proper support isn't provided, graduate student will keep struggling with the challenges mentioned, which will affect both the research they prepare and their chances to compete internationally.

Mastery of academic writing in English has become essential requirement for EFL learner. Especially their efforts have expanded beyond merely addressing grammar to include the ability to produce clear arguments with vocabulary and phrases aligned with global academic standards. As Aldabbus & Almansouri (2022) noted that many students spend time revising repeatedly their work simply to reach standards of coherence and accuracy. This challenge is common among Arab universities due to gab between Arabic rhetorical traditions and English academic discourse. Alzahrani (2025) added Students often struggle with cohesion, organization, and critical style due to limit practice and inconsistent feedback.

Contextual factors in Palestine, such as no access to writing support center, scarce resources, and reliance on lecture teaching, have all played a key role in reduce academic writing practice by students. Hammad (2016) highlighted that Palestinian students struggle into Structuring essay and including references properly, while Jabali (2018) stated that even enthusiastic learners experience frustration when attempting to create Simple and convincing writing. Such evidence underlines the importance of exploring new strategies, including the careful use of AI tools, to ease the burden of English research writing for graduate students.

1.2 Theoretical Background

This part included a review of the educational literature consisting of a theoretical framework and a set of previous studies related to the current study. The theoretical framework includes the educational literature that was referred to in this study.

1.2.1 The Emergence of Artificial Intelligence

Artificial intelligence is a branch of computer science that aims to create systems capable of performing tasks that usually require human intelligence, including learning from data, recognizing patterns, making decisions, and understanding natural language. AI encompasses a variety of techniques and methods that aim to make computers “smart” in the sense that they can mimic human mental abilities. The term artificial intelligence first appeared at Dartmouth College, in Hanover, USA, as a branch of computer science, in 1956 by John McCarthy. Vannevar Bush was the first scientist to propose a system that enables machines to think in the same way that humans do. This idea inspired computer scientist Alan Turing to study the possibility of simulating the human mind. Machines that could mimic human minds were developed and funded by the US Department of Defense, and researchers at the time predicted that in the future, machines would be able to perform tasks performed by humans (Alnafea, 2018).

There have been many achievements in the field of artificial intelligence, as the first mobile robot was created in 1969, but this invention was not widely accepted due to fears of negative effects. Attempts to use AI continued until the creation of a supercomputer, in 1997, which was able to defeat the world chess champion. Baidu launched the Linear Fold AI algorithm that was able to predict the DNA of a virus within 27 seconds. (Liu, Chen, & Yao, 2022).

1.2.2 The concept of artificial intelligence

The concept of artificial intelligence refers to computerized models applied in a variety of areas of life, through which basic relationships between elements are identified and appropriate reactions are selected in line with the given context (Yagazi, 2019). It also includes the ability to simulate human intelligence in decision-making through computer systems that are trained for this purpose (AlQarna, 2021). Moreover, artificial intelligence involves the ability to reach logical conclusions using technological information systems based on computers and other tools to complete tasks in a way that simulates the human mind (Coleman, 2020). In addition, it focuses on making computers intelligent and capable of thinking logically and solving issues through the use of mathematical methods similar to human thinking. AI also relies on gathering information from multiple sources and transforming it into useful knowledge, as well as choosing the appropriate algorithm

for the task at hand and continuously adjusting it to ensure accurate outcomes (Nitzberg & Zysman, 2021).

The researcher comes to the conclusion that artificial intelligence is the capacity of machines to make decisions that resemble human thought processes based on the information presented above.

1.2.3 Artificial Intelligence Characteristics

By imitating human thought processes, AI offers a set of technologies that enable computers to solve issues and form judgments just as effectively as people (Yagazi, 2019). Artificial intelligence is characterized by its capacity to learn autonomously from data and prior experiences without constant human support, as well as its capacity to adjust effectively to new and changing tasks and environments. In addition, AI can collaborate with humans in various fields such as logistics, healthcare, and manufacturing in order to enhance productivity and outcomes. It is also capable of intelligent interaction by communicating with users and offering helpful and informative replies to their needs and questions. In addition, AI can be accused of crimes such as reasoning and acting, as it depends on historical facts, data, analysis, and available information to make logical decisions. Another important characteristic of AI is its ability to handle massive amounts of data, as big data does not pose a challenge due to its capacity to analyze massive datasets at rapid speed and with limited time periods. Finally, AI continues to extend areas of use across a wide range of implementation, including medical diagnosis, data interpretation, prediction of behavioral patterns, and robot control, which increases its applicability and usefulness.

Artificial intelligence (AI) appears in several applications, including language learning and English language learning, where AI programs help students improve their language skills and effectively develop them. These applications provide customized programs for each student based on their English language level and provide interactive exercises that enhance their language skills (Sun, Anbarasan, & Kumar, 2020).

1.2.4 The Importance of Artificial Intelligence

According to the International Society for Technology in Education (2021) in its report *"Technology Learning for I"*, artificial intelligence plays an essential role in education due to its ability to create personalized learning experiences that align with students' needs and abilities. These experiences make learning more precise and efficient, as AI can also execute repetitive tasks that help maintain students' focus and keep them attentive to the learning process for longer periods (Zhang, 2022).

In this context, education is undergoing significant changes, and new educational tools such as machine learning and digital simulation are expected to assist learners in understanding more complex concepts and applying them as practical skills (Mu, 2019).

Moreover, AI applications enhance learners in various aspects by enabling self-assessment of performance, providing direct and consistent feedback, and offering interactive learning experiences. While these developments have the potential to transform education, it is essential to use AI technologies logically and ethically to ensure optimal and effective usage (Rad et al., 2018).

New technologies of the Fourth Industrial Revolution, such as artificial intelligence, big data, and 3D printing, play a vital role in many fields, including education and training. Thanks to these technologies, teaching and learning have fundamentally changed, freeing individuals from outdated practices and encouraging innovatio (Widiati, Rusdin, Indrawati, & Govender, 2023).

AI writing programs represent an advanced technology that enhances learners' writing skills. These programs are used in different ways, such as providing alternative words for spelling errors and offering additional context for suggested words and phrases. Despite the fact that the development of AI writing programs is still in its infancy, they provide useful support to writers in various areas and contribute to valid analysis and the development of writing skills (Alharbi, 2023)

1.2.5 Using Artificial Intelligence in Language Teaching

The use of technology, including innovative technology in the field of artificial intelligence, has become an integral part of foreign language learning and teaching in the current era (Rad et al., 2018). This role is to improve online foreign language learning

programs and mobile applications using AI technology. AI is used to enhance and assess learners' practices, guide them during the self-learning process, provide accurate evaluation of their performance, and correct their errors, as this method optimizes consciousness, develops abilities, limits learning costs, and enables learners to maintain information for an extended time (Huang , Zou, Cheng , & Chen, 2023).

In addition, programs like ProWritingAid help learners detect citation issues in essays, while apps like Grammarly help learners avoid grammatical errors and improve the quality of essays (Mainzaer, 2019). Chatbots can be an effective tool to provide help and answer questions for online learners, while automatic pronunciation training programs help learners improve authentic pronunciation and listening skills (Woo & Guo, 2023).

This holistic approach shows how AI technology is effectively improving the teaching and learning of foreign languages today (Pham, 2022). With the emergence of English as a global language, the ability to write well in English across diverse environments and for different audiences has become a necessity in second language education programs around the world. Teaching second language writing is often hampered by the amount of time and skill required to assess repeated drafts of students' writing, and automated online writing assessment programs have been developed as a way to address this challenge (Huang, et al., 2023).

In recent years, advances in artificial intelligence (AI) have greatly improved writing assessment and assistance systems based on automated writing assessment, text translation, synchronized feedback, and predictive text technology that automatically completes sentences. AI has gone further and can generate full texts by responding to a brief prompt from people (Alharbi, 2023).

Through advances in artificial intelligence based on a set of data that has become larger than ever before, a large number of AI applications have become available, each characterized by its interface, functionality, and target audience. It has also become possible to use a range of available smart writing tools to help learners improve the quality of their written texts, their acquisition of linguistic knowledge, and their ability to use technology (Nazari et al., 2021).

Pro Writing Aid is a comprehensive writing assistance application that seeks to assist writers enhance their writing skills and improve the quality of their written content. It offers a broad range of features that cover various aspects of the writing procedure, including grammar and spelling checking, where it helps writers detect and fix grammatical and spelling mistakes while offering clear guidance for revision. In addition, Pro Writing Aid improves writing style and readability by analysing the text and providing recommendations to improve clarity, expression, and overall readability. The application also includes a vocabulary enhancement feature that supplies synonyms and alternative phrasing to enrich vocabulary and avoid repetition. Moreover, Pro Writing Aid enhances report writing by generating detailed reports that point out areas for improvement and offer direction for better writing. It also contains a plagiarism checker that examines texts against academic and online databases to detect potential cases of plagiarism. Furthermore, Pro Writing Aid is characterized by its ease of integration and compatibility, as it can seamlessly integrate with various common writing platforms and software applications.

Pro Writing Aid is a powerful tool that helps students improve their writing skills and enhance the quality of their content. It also provides them with the ability to distinguish their level in the writing process and constantly try to improve it to reach the perfect text.

1.2.6 Graduate Education in Palestine

Graduate education in Palestine reflects both progress and challenges. On the one hand, universities in the West Bank, such as Birzeit, An-Najah National University, and the Arab American University, have expanded their postgraduate programs in recent years. Reports from the Palestinian Ministry of Higher Education show that about 1,900 students were enrolled in English-related postgraduate studies in 2023, a number that continues to grow despite political and economic difficulties (PCBS, 2023). On the other hand, the environment in which these students work is often unstable. Difficulties such as checkpoints, movement restrictions, and frequent electricity and internet outages have made it hard for students to attend classes or regularly access online resources (Khlaif & Salha, 2021).

In light of this, the quality of work submitted by graduate students is completely dependent on the resources they can access. Many students rely on online materials

because accessing international journals is restricted and expensive. For English language research, such restrictions must be taken into account, especially since students are required to meet international standards when writing academic work, while there are also few training opportunities available to students compared to those abroad. This is precisely what pushes students to postpone completing their theses (Hamdan, Ashour, & Daher, 2021).

Despite all the difficulties, higher education in Palestine is a model of resilience and defiance. Colleges and graduating students view education not only as a way to progress and achieve but also as a form of cultural revival for the nation. This is especially true in the field of English language, where developing graduate students' teaching is not merely an academic matter but an integral part of the effort to keep the voice of Palestine present on the world stage.

1.2.7 The Quality of Scientific Research in English

The research quality relies basically on how obvious and structured the writing is, because English is the language of nearly all significant journals internationally. UNESCO (2023) stated that approximately 80% of scientific papers are now published in English. This is not easy for students whose first language is not English, as they are expected to produce writing at the level of native speakers.

In the Arab countries, many graduate students still encounter difficulties when they have to write research in English. Studies in Jordan and the Gulf have shown that students often face challenges with paraphrasing, keeping arguments connected, and using academic style correctly (Shuqair & Suleiman, 2023). Regarding Palestinian universities, the situation is considered tough: supervisors want their students to publish and compete internationally, while they do not get enough training to master writing skills. A study by Hamdan, Ashour, & Daher (2021) discovered that Palestinian postgraduate students' research is weak and not readily accepted globally since they have issues with their writing, such as phrase repetition, literal translation, and limited vocabulary.

The issue is not only about students' language but also about resources. The Palestinian universities' libraries are out of date, and internet connection is unstable, so students cannot always access updated studies. This means that even when students know what to

write, they may not have the material to support their ideas. It is common for students to take much longer to finish their theses compared to other countries.

In recent years, some of these challenges have been reduced by technology. Writing tools, including AI-based applications, are now helping students check grammar, improve style, and organize references. Liu & Wang (2025) explain that such tools are useful, but they should not replace critical thinking. The best research comes when students combine their own effort and originality with the support of technology.

For Palestinian graduate students, improving English writing is not just a skill for passing courses. It is the key to producing research that can stand next to work from other countries and make the Palestinian academic voice heard in the world.

1.2.8 Domains of the Study

This study is arranged based on three main domains that introduce the ways in which artificial intelligence tools affect English academic research. The first domain investigates the effect of AI on the quality and efficiency of research, looking at how these tools can improve writing quality. The second domain points out the strengths and weaknesses of AI tools. Although students may benefit from using these tools in grammar editing and time optimization, they may also encounter problems such as providing incorrect phrases or a lack of originality. The third domain introduces strategies to achieve optimal usage of AI tools, such as the balanced method adaptation to combine human critical thinking with AI tools' assistance.

First Domain: The Effect of AI on Research Quality and Efficiency

Artificial intelligence plays a vital role in enhancing the quality of academic research, especially in English writing. For Palestinian graduate students, these tools assist students by providing the necessary support in fields that are often considered the most challenging, such as paraphrasing, formulating, and data organization.

Recent research indicates that AI helps to improve research quality in various aspects. It improves writing accuracy, as students can better meet writing standards by using applications like Grammarly or Pro Writing Aid. These applications minimize errors and enhance accuracy. AI also helps save time by optimizing the research writing process by providing corrections and creating references, which allows students to concentrate more

on the content. In addition, AI tools enhance data analysis by enabling students to handle large datasets successfully, which enhances the reliability of their results. Moreover, AI tools support organization by helping students structure arguments and maintain coherence in research.

At the same time, AI's role in efficiency is relevant to the Palestinian context, where students face challenges in accessing resources. The ability to perform tasks quickly gives students the chance to spend more time focusing on originality and critical thinking. However, as referred to in recent studies, AI tools are most effective when they are used to support, rather than replace, the student's personal effort (Fitria, 2021 b).

Second Domain: Strengths and Weaknesses of AI Tools

Artificial intelligence tools have become integrated part of higher education. For many Palestinian, Adopting its use is no longer optional but a demands to conduct English research. However, AI's tools offer benefits to be gained and drawbacks to be avoided.

AI tools demonstrate several strengths in supporting graduate students' academic writing. One of the most important strengths is language accuracy and clarity, as AI has the potential to deveop linguistic accuracy, which is considered one of the most obvious benefits for graduate students. Tools such as Grammarly and Pro Writing Aid help address challenges related to grammar and organisation structure, and Chien (2023) noted that the use of AI for language checking supports text clarity. For Palestinian students in particular, AI tools make presenting complex ideas in English easier and less demanding. Another key strength is time efficiency, as academic writing needs tremendous effort to refine language and style. Xu (2025) reported that approximately 50% of postgraduate students used AI tools to improve text readability, while 22% relied on them for grammar correction, reinforcing the idea that AI functions as an effective time keeper. In the Palestinian context, where academic resources are restricted, AI tools help save time and minimise effort while maximising benefit. In addition, AI provides strong assistance for organisation and coherence, as its role extends beyond spelling and grammar to organizing arguments and preserve coherence in lengthy academic works. Kim et al. (2023) found that these capabilities director students effectively, particularly in major projects such as theses and dissertations. Finally, AI tools support accessibility in restricted -resource contexts, as Palestinian university students have serious challenges

related to lack of resources and unstable internet access (Smith & Scott, 2023). Even when access is limited, AI tools can reduce academic support gaps by providing instant feedback, enabling students to improve their writing skills independently without constant outside guidance.

Despite its strengths, artificial intelligence has several weaknesses that limit its reliability in academic work. One major weakness is the danger of incorrect data, as AI sometimes presents incorrect citations and inaccurate facts. Sun, Gregory, & Polastri (2024) emphasized a repeated problem in language models known as "hallucination," which involving inaccurate research findings and, in some cases, generating citations that do not exist. For graduate students, such errors can seriously damage their work credibility. Another weakness relates to cultural and contextual insensitivity. Although AI demonstrates effective performance in correcting grammar, it often fails to produce text that is contextually grounded. While the language may be grammatically correct, it can appear unnatural and lack the subtlety needed in academic writing. The Stanford Review (2024) identified this challenge as especially problematic for EFL writers, who may depend on AI and consequently produce text that lacks their original voice. In addition, heavy reliance on AI can limited critical thinking, as relying entirely on AI is not recommended and committing errors remains an important part of the learning development for students. Chen & Li (2023) emphasized that AI should be viewed not as a substitute but as a supportive partner, since creativity and critical engagement stay essential and cannot be reduced or substituted by AI. Finally, ethical and integrity concerns arise from the use of some rewriting and summarization tools, which may produce text that is very similar to original sources, raising concerns about plagiarism and academic ethics. Musa and Almryad (2025) warned that careless use of such tools may reduce academic credibility, even though they can assist in rephrasing text.

Third Domain: Strategies to Maximize the Benefits of AI Tools

The use of AI in postgraduate research is still fresh, and many students don't have enough experience in this domain so taking strategies into consideration is a necessity. In Palestine, this issue becomes even more urgent because students face real obstacles such as limited resources, unstable internet access, and a lack of specialized training in academic writing. The results of the study showed that students had a generally positive

perspective about strategies that could assist student to get benefit from AI tools, but they also recognized that the tools should not taking place their own effort.

One prominent strategy is choosing the right tool for the right task. With so many AI applications on the table, students need to choose carefully. A grammar checker may be useful in editing, while a reference manager may be more suitable for the final stage of writing. As Liu & Wang (2025) explained, the best outcomes come when students match the tool with the research stage instead of using one system for everything.

Another useful strategy is integrated AI with traditional methods. Mind mapping, peer work, and teacher's role are still important. Chen & Huang (2024) noted that students achieved the greatest development when AI was used Along with human feedback. This balance is very important in the Palestinian situation, where shared learning and instructional guidance play important role in high education.

A third strategy is checking AI outputs. Tools produce errors and often invent false responses. Gutiérrez's (2023) viewed that students who verified AI suggestions with academic guidelines and library databases created stronger work. This increases the authenticity of the research and decrease the spread of wrong information.

Moreover, students also in need to improve critical consciousness. Ju (2023) stated that over dependence on AI without reflection decreased originality by more than 20%. Students have to keep their voice obvious instead of just copying the AI suggestions.. In Palestinian universities, short workshops on AI literacy in doing research courses could help students increase this awareness.

Finally, there is the issue of ethics and academic accuracy. The increased use of paraphrasing or content-generation tools may create problems with plagiarism. Alhajji (2024) suggested that universities should put obvious rules on how AI can be used and enhance these rules with awareness campaigns. For Palestinian institutions, adding AI ethics into research regulations could ensure that students benefit from technology while protecting academic honesty.

As the researcher notes that these strategies point out that AI do best if students deal with it as supportive tool, not as a replacement. When students choose tools with care, merge them with their original practices, check outputs, think critically, and provide

ethical rules, they can get the optimal usage of AI without losing their autonomy or individuality.

Conclusion of the Three Domains

The three domains show a picture of how AI tools are contributing in graduate research in Palestine. The first domain indicated that students perceive AI to have an impact on the quality of their academic writing, especially in grammar, vocabulary, and research structure. This suggests that technology can decrease some of the language challenges that many students encounter when writing in English.

The second domain highlighted the strengths and weaknesses of AI tools. AI is a double-edged sword: on one hand, it reduces time, fixes mistakes, and creates clear texts; on the other hand, it may provide fake data, neglect cultural context, and sometimes reduce students' autonomy if used without evaluation. These results draw attention that AI tools can support or harm research based on how they are used.

The third domain addressed strategies to achieve optimal use of AI tools. Students emphasized the importance of deciding on proper tools, integrating AI with traditional practices, reviewing outputs, and keeping ethical criteria. Overall, the three domains stressed that AI can play a key role in academic research in Palestine, but only when used properly.

Foreign Studies

Kurniati & Fithriani (2022) examined how graduate students engage with QuillBot in their writing. They observed that students described the tool as both easy and helpful. It did not replace their effort but made the process easier by offering fast revision and concise expression. Remarkably, many participants noted that QuillBot increased their confidence, as they sensed more able to create texts that fit academic standards. From my point of view as a researcher, this highlights a significant consideration that AI tools can offer more than just fixing language mistakes; they can make students feel comfortable.

In another context, Zhang & Wu (2024) collaborated with science and technology postgraduate students to enhance an academic writing course. Their method combined cooperative learning, awareness of writing form, and digital AI assistance. Students benefited from having directed support during various steps of writing, from

brainstorming ideas to polishing the final draft. Yet, the study also showed a note of carefulness. Some learners depended too much on AI, and they are in danger of a decline in their own essential involvement. For this study, that tension is relevant to the Palestinian context, since Palestinian graduate students are exposed to the same risks related to over-dependence on AI when other academic resources are not available.

Smith & Scott (2023) provided a clearer picture through their overview of AI use in higher education. They asserted the increasing role of AI across writing, reading, and autonomous learning. At the same time, they real-life challenges such as technical issues and the danger of standardizing language in ways that dismiss local diversity. For me, the value of this review lies in highlighting that AI is not neutral. It offers support, but it also changes academic processes in ways that may not suit every context. From my point of view, this represents pressure in itself when discussing the situation in Palestine, where access to the internet and important international journals is relatively restricted compared to more advanced regions globally.

1.3 Arab Studies

Research in Arab countries has also stressed the contribution of AI in language education, although most of it has concentrated on primary and secondary education, not on postgraduate research writing. Mohammed (2024), for example, experimented with AI applications in teaching English to primary school pupils in Aleppo. The findings were surprising. Students in the experimental group excelled better in language skills than their peers. While this context is away from graduate-level writing, the study does confirm that AI can make an important change in how learners acquire and practice English. This enhances the idea that AI could support advanced learners in academic tasks if they apply it appropriately.

Al-Hakami (2023) looked at AI in Saudi Arabia’s public education system. His study focused on how political and economic factors have affected the adoption of AI in schools. Although the findings presented optimism about the role of AI in enhancing learning, they also pointed to important challenges. I find this viewpoint very important because it alerts us that AI is never used in isolation. Social and institutional factors create both its opportunities and its limits. Similarly, in Palestine, incorporating AI into

university research writing is affected by the same constraints related to the lack of resources and political challenges.

Finally, Abdulwahab & Colleagues (2023) examined the effect of AI tools on high-achieving secondary students in Egypt. Their findings highlighted important development in creative linguistic skills, with a clearly large effect size. This affirms that AI can do more than correct language; it can also foster innovation. For graduate students who have difficulties with English academic writing, this possibility should get attention. If AI can assist younger learners reflect flexibly in another language, it can also improve advanced students in improving arguments and structuring research more efficiently.

1.4 Local studies

Through reviewing local studies on the use of Al tools in Palestinian higher education, it is clear that AI serves a supportive role in improving the quality of academic writing in English.

A local study examined by Yousef (2025) investigated the use of artificial intelligence tools among medical students in Palestinian universities. The findings showed that AI tools were widely used to assist academic and research writing in English, particularly in enhancing language accuracy, clarity, and organization of research texts. Students stated that AI-supported tools assist them reduce linguistic difficulties, develop ideas, and revise their academic work more productively. However, the study also highlighted concerns regarding heavy reliance on AI and highlighted the importance of supervised and responsible use. These results indicate that AI can positively enhance to improving the quality of English-language research writing within the Palestinian higher education context, which is consistent with the focus of the present study.

Another local study conducted by Hamayil and Ghayyadah (2025) investigated the role of generative AI tools in supporting the educational experience and learning outcomes among postgraduate students at Al-Quds Open University. The findings showed that these tools facilitated access to information and enhanced learning processes, contributing to improved educational outcomes. However, the study also highlighted serious challenges associated with integrating generative AI, including the need for training and institutional frameworks to guide its effective use. In my view as a researcher, I believe these studies emphasize the capacity of AI to assist research writing, however they also indicate notable

deficiencies, such as heavy reliance on the tools and limited emphasis on building comprehension of research methods. Therefore, I argue that incorporating AI into academic research demands structured and ethical use, together with clear policies and guidelines, to guarantee the efficient and equitable improvement of research quality within the Palestinian context.

Based on the findings of (Yousef, 2025; Hamayil & Ghayyadah, 2025; Qashoo, 2025) offers further understanding by through exploring practical experiences with AI tools at Al-Quds Open University. While Yousef emphasised the beneficial impact of AI on English research writing quality and Hamayil and Ghayyadah stressed enhancement in learning outcomes, Qashoo shows that students face real challenges, difficulties especially concerning proper training and institutional support. From my view as a researcher, this indicates that although AI has evident potential to improve academic performance and research writing, its effectiveness relies strongly on organized application, adequate support, and students' competence. Therefore, incorporating AI in Palestinian higher education demands not only the accessibility of tools but also directed, responsible use together with explicit rules and instructional frameworks.

In brief, local studies emphasize that AI tools can improve scholarly writing, facilitate learning outcomes, and help learners in research tasks. Their efficiency, nevertheless, relies on structured use and structured assistance, forming a basis for the present study and comparison with Arab and international research.

1.5 Statement of the Problem

In recent time, the improvement of artificial intelligence (AI) has transformed higher education and academic research. For Palestinian graduate students, who face complicated difficulties in creating high-quality research in English, AI tools could offer the necessary assistance. However, the condition is still ambiguous because of a lack of evidence regarding their effect, especially in a place like Palestine where resources are restricted and the academic context is not well-prepared.

This research is significant because it underscores a clear deficiency: there is a limited local evidence investigating the effect of AI tools on writing quality in the Palestinian context. This issue is important, as students encounter difficulties with academic writing, restricted resources that support academic writing, and political and infrastructural

constraints. Therefore, this study examines the perspectives of graduate students regarding whether AI tools have a positive impact on research quality in the Palestinian situation.

Many Palestinian students struggle with difficulties that make it hard to acquire English for academic practice. They suffer from restrictions such as limited reach to relevant resources, political constraints, and poor infrastructure. Studies have stated that Palestinian students have difficulties with speaking and writing in English, due to stress, few chances to practice, and curricula that don't always give priority to communicative skills (Stock & Ron, 2023). Inability to incorporate modern technology into teaching is also a main obstacle due to inconsistent infrastructure, such as limited internet access or outdated devices, which makes implementing AI even harder (Alshamy, Al-Harthi, & Abdullah, 2025).

To explore this issue locally, preliminary informal interviews were conducted with English instructors in Jenin as part of my preliminary research. Most of them said that they rarely implement AI in their classes because of the insufficient technical assistance and professional development, but they all agreed with the use of AI in the classroom. They assured that it could make a major improvement because students are tech-savvy, and AI addressed students' needs, contributing to progress in teaching writing.

To see the situation from a broader view, research from other contexts suggests AI can change academic writing. For example, tools like AI writing systems have been found to optimize time in writing while raising quality from average to excellent, at least in controlled studies (Marzuki, Widiati, Rusdin, Darwin, & Indrawati, 2023). Yet, there is another side: heavy dependence on AI can lead to decrease creativity or even ethical challenges, such as unintended plagiarism (Einarsson, Lund, & Jónsdóttir, 2024). In Palestine specifically, the application of AI tools in higher education is amidst ongoing challenges; comprehension this is important for improving students' writing outcomes.

This study intends to address the gab by focusing on how AI impacts the quality of English research as perceived by Palestinian graduate students. The study examines what they know about these tools, the advantages and disadvantages, and ways to use them correctly. Therefore, the main research question is:

What is the effect of AI on the quality of English research, as perceived by Palestinian graduate students?

1.6 Significance of the Study

First: Theoretical Importance

This study adds to theoretical knowledge in the domain of using artificial intelligence applications in English language research, particularly in enriching the overall quality of graduate students' academic writing. It provides new understanding of the role of AI in supporting the research writing process and motivates graduate students to implement AI as a supportive tool in academic research, which enhances research outcomes while maintaining quality. In addition, the study highlights the significance of keeping up with the demands of integrating the digital age into higher education. The findings also help in supplying practical frameworks for developing postgraduate research, especially in situations where students encounter challenges in academic writing and research structuring. Furthermore, this study contributes to Arab libraries by providing new research that investigates the use of artificial intelligence in supporting the quality of academic research written in English.

Second: Applied Relevance

The study presents recommendations that can enhance the improvement of graduate curricula in Palestine by emphasizing the significance of keeping up with the digital age and addressing students' needs. It provides evidence regarding the benefits of artificial intelligence and its role in achieving the purposes of graduate research, particularly under limitations related to the lack of resources and access. In addition, the findings can support Palestinian universities in integrating AI training into research courses, providing students with opportunities to use these tools effectively. Furthermore, the study encourages further research in the field of AI and its role in raising the quality of academic outputs written in English.

1.7 Aims of the Study

The current study aims:

1. To evaluate the perceived efficiency of AI among Palestinian graduate students' research.
2. To diagnose the benefits and challenges faced by Palestinian graduate students when using AI tools to improve their research writing skills.

3. To propose AI-based strategies that can support Palestinian graduate students in achieving higher competence in English research writing.

1.8 Questions of the Study

1. How efficient are AI tools in enhancing the quality of English language research as perceived by Palestinian graduate students?
2. What are the benefits and challenges faced by Palestinian graduate students when using AI tools to improve their research writing skills?
3. What AI-based strategies can be implemented to support Palestinian graduate students in achieving higher competence in English research writing?

1.9 Limitations of the Study

This study has certain limitations that require further studies by other researchers, including:

1. Time limit: This study is conducted in the academic year 2024 - 2025.
2. Place limit: This study is conducted in the Palestinian state.
3. Topic limit: The topic itself is 'The effect of AI on the English language research quality of Palestinian graduate students'.

1.10 Hypotheses of the Study

1. There are no statistically significant differences in the perceived effect of AI tools on the quality of English language research among Palestinian graduate students according to the gender variable.
2. There are no statistically significant differences in the perceived effect of AI tools on the quality of English language research among Palestinian graduate students according to the age variable.
3. There are no statistically significant differences in the perceived effect of AI tools on the quality of English language research among Palestinian graduate students according to the IT knowledge variable.

4. There are no statistically significant differences in the perceived effect of AI tools on the quality of English language research among Palestinian graduate students according to the AI experience variable.

Chapter Two

Methodology

2.1 Study design

A descriptive, quantitative technique was employed since it matched the study objectives. The researcher employed a questionnaire to acquire study data. The effect of AI on English language research quality as perceived by Palestinian graduate students. The researcher used a questionnaire as the main tool for the study.

2.2 The Study Population and Sample

The study's population was consisted of 1,931 (graduate students joined in English language programs) in the West Bank. This population was chosen because the research is concerned with how AI tools affect the quality of students' academic research.

According to statistical guidelines, a sample of about 320 students would normally be recommended for a population of this size in order to achieve a 95% confidence level with a 5% margin of error (Sample Size Essentials for Educational Studies, 2025). This number was therefore set as the target sample.

Out of the targeted 320 participants, a total of 171 valid responses were collected. The smaller number was mainly due to the inability to reach all students because of Gaza ware and mobility restriction. While this figure is lower than the recommended sample size, it still represents a meaningful portion of the population (about 8.9% of the total). More importantly, the retrieved responses came from a variety of students—male and female, from different age groups, and with varying levels of IT knowledge and AI experience. This diversity adds balance to the data and helps ensure that the findings reflect a wide range of perspectives within the student population.

Although the initially targeted a sample of 320, only 171 were collected because the circumstances that affected reach to all students .The sample is less than what expected but still represent significant portion of population (about 8.9% of the total).The background of the study sample is varied in term of gender, age groups, IT knowledge and AI experience, as shown in Table 1 bellow.

Table (1)

The study sample distribution by gender, age, IT knowledge, and experience using AI

Variables	Frequency	Percentage
Distribution of Sample According to Gender		
Male	33	19.3%
Female	138	80.7%
Total	171	100%
Distribution of Sample According to Age		
22-30 years	80	46.8%
31- 39 years	68	39.8%
40 years and more	23	13.5%
Total	171	100%
Distribution of Sample According to Having IT Knowledge		
Low	10	5.9%
Medium	79	46.2%
High	82	48.0%
Total	171	100%
Distribution of Sample According to Having Experience in		
Low	28	16.5%
Medium	77	45.3%
High	65	38.2%
Total	171	100%

2.3 Instrument of the Study

To achieve the objectives of the study, the researchers used a 24-item questionnaire for graduate student who enrolled in English program. The questionnaire had two main sections. The first section focused on demographic data such as gender, age, level of IT knowledge, and experience with AI. The second consisted of three domains of Assessing the effect of AI tools on research quality and efficiency, Identifying the strengths and weaknesses of AI tools, and Strategies to maximize the benefits of AI tools. The researcher used a five-point Likert scale that includes strongly agree, agree, neutral, disagree, and strongly disagree.

2.4 Validity and Reliability of the Instrument

To guarantee the validity of the questionnaire, it was evaluated by a jury of experts in the field of teaching. The jury of experts consisted of three professors teaching at An-Najah National University and the Arab American University, and the last expert is specialized in AI applications. A total of four experts, including Dr. Zuhair Khlaief, Dr. Khalid Dwaikat, Dr. Nizzar Millhem, and Mr. Mohammad Khamaiseh, examined the questionnaire. The jury's recommendation was considered to revise and enhance the questionnaire's content. Their assessment focused on the clarity, accuracy, and appropriateness of the items in relation to the research aims. Based on their guidance, items were revised. The validity of the questionnaire was achieved after completing the modifications, and the final version of the questionnaire was in its final form.

Furthermore, the questionnaire's reliability was assessed using Cronbach's Alpha Equation.

Table (2)

Reliability coefficients of each domain and the total score of the study

Domains	No. of items	Reliability coefficient
Assessing the Effect of AI Tools on Research Quality and Efficiency	8	0.747
Identifying Strengths and Weaknesses of AI Tools	8	0.630
Strategies to Maximize the Benefits of AI Tools	8	0.865
Total score	24	0.907

2.5 Procedures of the Study

Data collection

The final draft of the questionnaires was given to study sample online. It took about 6 weeks for the instruments to be distributed, collected, and returned to the researcher. A total number of valid response is 171.

The research is based on two basic sources of data:

1. Primary: Data This study used a questionnaire to assess the effect of AI on English language research. The structured questionnaires dustributed electronacly. it took

about 6 weeks for instrument to be distributed, collected and returned to the researcher. the it is analyzed statistically using the SPSS computer software.

2. Secondary: data was collected from books, journals, papers, and websites to examain the context. A literature review was help to prepare the study and identify its objectives, with a focus on the effect of AI on English language research quality.

The questionnaire conssist of two main sections:

1. Part one: This section aimed to collect respondent demographics, including gender, age, AI experience and IT Knowledge
2. Part two: This section included 24 items.researcher used a five-point Likert scale (5 = strongly agree, 4 = agree, 3 = neutral, 2 = disagree, 1 = strongly disagree). The items were consisted of three domains: the effect of AI on English language research quality, strengths and weaknesses of AI tools, and Strategies to maximize the benefits of AI tools.

This questionnaire was administered online because the actual process of data gathering encountered significant difficulties. Because of the continuing war in Gaza, universities there were excluded. Even within the West Bank, mobility restrictions and the unstable situation made it difficult to reach students directly on campus. Because of these conditions, the researcher depended on an electronic questionnaire, which was distributed through the Deanship of Student Affairs at An-Najah National University to guarantee that graduate students across different universities could still be reached. Johnson, Lee, & Martinez (2024) confirmed that employing electronic surveys has been recognized as efficient and feasible way of gathering data during crises.

In summary, this chapter introduced the way researcher follow to colloct data and the study design used to conduct this study and methods were used for analysis. All of these help researcher to present the finding in the next chapter.

Chapter Three

Results

3.1 Introduction

This chapter aims to identify the effect of AI on English language research quality for graduate Palestinian students. The collected data completed by graduate students across universities in the West Bank during the academic year 2024–2025. It also aims to identify the effect of several variables on the responses. To accomplish the aims of the study, the researcher analyzed the data in accordance with the study questions, and the results were as follows:

3.2 Results related to the study questions

3.2.1 Results Related to the main question

The main question of this study was:

"What is the effect of AI on the quality of English research, as perceived by Palestinian graduate students?"

To answer this question, the researcher analyzed the responses by calculating the means, standard deviations, and percentages of each item. Furthermore, the researcher used the following estimates to differentiate the degree of approval:

1. The arithmetic mean (4.21 or more, equivalent to 84.2% or more) has a very large degree.
2. The mean (3.41 - 4.20 and 68.2% - 84.0%) has a high degree.
3. The mean (2.61 – 3.40 and 52.2% - 68.0%) has a medium degree.
4. The mean (1.81 – 2.60 and 36.2% - 52.0%) has a low degree.
5. The mean (less than 1.81) has a very low degree.

It presents the ranges of arithmetic means and percentages that indicate whether the effect of AI tools is viewed as very high, high, medium, low, or very low. This classification helps in understanding how strongly the participants agreed with the impact of AI tools on their research quality (Doronila, 2025).

Table (3) demonstrates the results:

Table (3)

The applied statistical tests on the effect of AI on the English language research quality of Palestinian graduate students

No.	Domain	M	SD	%	Level
1	Assessing the Effect of AI Tools on Research Quality and Efficiency	3.595	0.676	71.91%	High
2	Identifying Strengths and Weaknesses of AI Tools	3.521	0.729	70.43%	High
3	Strategies to Maximize the Benefits of AI Tools	3.620	0.633	72.39%	High
Total Scores		3.579	0.682	71.58%	High

Table (3) presents the averages and standard deviations for the three study domains. In the first domain, which discusses the effect of AI tools on research quality and efficiency, the average score is relatively high. This suggests that many participants believe that AI tools are beneficial in making their research practices easier and more effective.

The second domain relates to the strengths and weaknesses of AI tools. The mean is also in the high range. This means that participants see clear benefits from using AI tools, but at the same time they are aware of some problems or limits.

The third domain, which deals with strategies for getting the most out of AI tools, has the highest mean. This shows that participants believe strategies and practical steps are very important when using AI.

When looking at the total score for all domains, the result is also high. This indicates that, overall, participants have a positive view of AI tools in research, even though they recognize that these tools are not perfect.

3.2.2 Results related to the sub-questions

3.2.2.1 Results related to the first sub-question

How efficient are AI tools in enhancing the quality of English language research as perceived by Palestinian graduate students?

The efficiency of AI on Palestinian graduate students' research.

Table (4) demonstrates the outcomes:

Table (4)

The applied statistical test on the efficiency of AI on Palestinian graduate students' research

No.	Item	*M*	*SD*	*%*	Level
1	AI tools improve the quality of my academic research.	3.720	0.528	74.39%	High
2	Using AI tools saves my time during the scientific research process.	3.882	0.353	77.65%	High
3	AI tools help me achieve more accurate results in my research.	3.537	0.657	70.73%	High
4	AI tools make it easier for me to collect data.	3.587	0.710	71.74%	High
5	AI tools enhance the organization and structure of my research.	3.733	0.561	74.67%	High
6	AI tools make it easier for me to analyze data.	3.515	0.688	70.30%	High
7	The cost of accessing advanced AI tools limits my ability to use them effectively in my research.	3.420	0.799	68.40%	High
8	I find it difficult to access suitable AI tools for my research needs.	3.267	0.745	65.34%	Medium
Total score		3.595	0.676	71.91%	High

Table (4) presents the mean and standard deviation for the eight items that make up the first domain. The results show that most items have high mean scores, which indicates that participants agree on the positive role of AI tools in their research. For example, the items about saving time, improving the quality of research, and supporting data analysis all received high scores. These findings introduce the idea about how participants

perceive AI tools. They view it as practical and beneficial tool in different field of their academic work.

Simultaneously, two items linked to cost and access scored lower than the others. This means that some participants believe that they suffer from financial limits or difficulties in having the prober AI tools that match their needs. While these concerns are important, they do not reduce the overall positive view.

In general, the results of Table (4) suggest that participants believe AI tools support their research by improving quality, saving time, and organizing information, although cost and access remain as challenges.

3.2.2.2 Results related to the second sub-question

The benefits and challenges faced by Palestinian graduate students when using AI tools to improve their research writing skills.

Table (5) demonstrates the outcomes:

Table (5)

The statistical tests on the benefits and challenges faced by Palestinian graduate students when using AI tools to improve their research writing skills

No.	Item	*M*	*SD*	*%*	Level
1	Recall (Strength) AI tools like Grammarly and Turnitin help me correct grammar and spelling errors in my academic writing.	3.615	0.651	72.29%	High
2	Comprehension (Strength) Using AI tools improves the clarity and structure of my sentences in English research papers.	3.712	0.511	74.23%	High
3	Application (Strength) AI tools save my time by automating repetitive tasks like proofreading and formatting references.	3.697	0.536	73.94%	High
4	Synthesis (Strength) AI tools provide suggestions for improving the vocabulary and terminology used in my research.	3.650	0.634	72.99%	High
5	Analysis (Weakness) AI tools sometimes fail to understand the cultural or academic context of my research topic.	3.529	0.720	70.58%	High
6	Evaluation (Weakness) AI-generated content often lacks originality and depth compared to human-written text.	3.362	0.794	67.24%	Medium
7	Application (Weakness) I find it difficult to rely on AI tools for complex tasks like analyzing data or interpreting results.	3.146	0.879	62.92%	Medium
8	Analysis (Weakness) AI tools can produce inaccurate or irrelevant suggestions that require manual correction.	3.462	0.779	69.23%	High
Total score		3.521	0.729	70.43%	High

Table (5) shows the results for the second domain, which handles with both strengths and weaknesses of AI tools. Several items associated with strengths received higher ratings. For example, participants approved that AI helps with grammar correction, produce clear sentences, and minimize time in doing their tasks such as proofreading and formatting. They also noted that AI tools can provide better vocabulary, something that many found useful for academic writing.

On the other hand, the items represented weaknesses had lower scores in comparison with the strengths. Some participants refers that AI doesn't often succeed in understanding the research context, especially cultural or academic issues. Others pointed out that AI text sometimes seems less original and less deep than human writing. Additionally, several participants indicated that relying on AI for complex tasks like data analysis or interpretation sometimes produces inaccurate or irrelevant output.

Taken together, the results suggest that participants see AI tools as very helpful for simple or routine work, but they are more cautious when it comes to deeper or more complex research tasks. In general, they treat AI as a support tool rather than a complete replacement for human effort.

3.2.2.3 Results related to the third sub-question

What AI-based strategies can be implemented to support Palestinian graduate students in achieving higher competence in English research writing?

The results are displayed in Table (6):

Table (6)

The statistical tests on the AI-based strategies can be implemented to support Palestinian graduate students in achieving higher competence in English research writing

No.	Item	M	SD	%	Level
1	Participating in training sessions on AI tools would enhance my ability to effectively utilize these tools for improving the quality of my English language research.	3.645	0.584	72.91%	High
2	Collaborating with peers or professors helps me use AI tools more effectively.	3.555	0.672	71.09%	High
3	Having a user guide for AI tools would be beneficial for my research.	3.603	0.631	72.07%	High
4	Selecting AI tools that match my research topic improves outcomes.	3.720	0.568	74.41%	High
5	Using AI tools strategically at different stages of research (e.g., planning, drafting, revising) helps me maximize their benefits without becoming overly dependent on them.	3.719	0.581	74.38%	High
6	Using AI tools alongside traditional research methods (e.g., brainstorming, peer feedback) helps me maintain my critical thinking and creativity in research.	3.680	0.579	73.61%	High
7	Implementing strategies like cross-checking AI outputs with academic guidelines helps me overcome ethical concerns about AI-generated content.	3.562	0.674	71.24%	High
8	The lack of transparency in how AI tools generate results undermines their reliability for my research.	3.471	0.714	69.43%	High
Total score		3.620	0.633	72.39%	High

Table (6) presents the results for the third domain, which focuses on strategies for making the best use of AI tools. Most of the items in this table scored high, showing that participants see clear value in planning how AI is used during research. For example, they

agreed that training sessions and user guides would help them use AI more effectively. They also believed that working with peers or professors improves their ability to benefit from these tools.

Another significant point in the table is the use of AI in various stages of research. Acoording to Participants, Applying AI tools to plan, draft, or revise can help them with making the process more effective. they also confirmed about the importance of not relaying on it. The results bought attention to the strategy of integrating AI with traditional research practices.it shows that this strategy was perceived positively.

At the same time, several concerns were noted. Some students stated that the lack of transparency in how AI produces results could decrease trust in it. Still, the overall picture is clear: participants support the idea of using strategies to maximize the benefits of AI, while keeping a balance between technology and critical human judgment.

3.2.3.1 Results related to the hypotheses

There are no statistically significant differences in the perceived effect of AI tools on research quality according to the three variables.

Table (7)

Summary of Hypothesis Tests by Grouping Variables

Grouping variable	Domain	Test	*Df*	Statistic	p	Conclusion	Effect size
Gender	Impact on quality/efficiency	T	169	$t = 0.10$	$p = .92$	Not significant	$d \approx 0.02$
Gender	Strengths/weaknesses	T	169	$t = 0.15$	$p = .88$	Not significant	$d \approx 0.03$
Gender	Strategies	T	169	$t = 0.90$	$p = .37$	Not significant	$d \approx 0.17$
Gender	Total score	T	169	$t = 0.75$	$p = .45$	Not significant	$d \approx 0.15$
Age	Impact on quality/efficiency	ANOVA	2,149	$F = 1.45$	$p = .238$	Not significant	$\eta^2 \approx .019$
Age	Strengths/weaknesses	ANOVA	2,158	$F = 1.73$	$p = .180$	Not significant	$\eta^2 \approx .021$
Age	Strategies	ANOVA	2,156	$F = 1.03$	$p = .361$	Not significant	$\eta^2 \approx .013$
Age	Total score	ANOVA	2,166	$F = 1.19$	$p = .306$	Not significant	$\eta^2 \approx .014$
IT knowledge	Impact on quality/efficiency	ANOVA	2,149	$F = 0.73$	$p = .482$	Not significant	$\eta^2 \approx .010$
IT knowledge	Strengths/weaknesses	ANOVA	2,158	$F = 1.66$	$p = .193$	Not significant	$\eta^2 \approx .021$
IT knowledge	Strategies	ANOVA	2,156	$F = 6.77$	$p = .002$	Significant	$\eta^2 \approx .080$
IT knowledge	Total score	ANOVA	2,166	$F = 2.72$	$p = .069$	Marginal / trend only	$\eta^2 \approx .032$
AI competence	Impact on quality/efficiency	ANOVA	2,148	$F = 1.42$	$p = .246$	Not significant	$\eta^2 \approx .019$
AI competence	Strengths/weaknesses	ANOVA	2,157	$F = 0.76$	$p = .470$	Not significant	$\eta^2 \approx .010$
AI competence	Strategies	ANOVA	2,155	$F = 5.47$	$p = .005$	Significant	$\eta^2 \approx .066$
AI competence	Total score	ANOVA	2,165	$F = 2.73$	$p = .068$	Marginal / trend only	$\eta^2 \approx .032$

Note. df = degrees of freedom. All p values are two-tailed. Effect sizes are Cohen's d (t-tests) or η^2 (ANOVAs). Significant rows are bolded.

Table (7) documents the results of the hypothesis tests concerning the perceived effect of AI tools on the quality of English-language research among Palestinian graduate students across the classification variables (gender, age, and IT knowledge). The table reports the test type, degrees of freedom (*df*), test statistic (*t*/*F*), two-tailed *p* value, and effect size (Cohen's *d* and η^2) at ($\alpha = 0.05$).

Adjudication of the three hypotheses based on Table 7 (quality/efficiency domain).

Hypothesis 1 (Gender). An independent-samples T-test yielded t(169) = 0.10, p =.92, with a very small effect size (d ≈.02). There is no statistically significant difference in the perceived effect by gender; the null hypothesis is retained, and practical significance is negligible.

Hypothesis 2 (Age). A one-way ANOVA indicated F(2,149) = 1.45, p =.238, η^2 ≈.019 (small). There is no main effect of age on the perceived effect; the null hypothesis is retained, and no post-hoc comparisons are warranted.

Hypothesis 3 (IT knowledge). A one-way ANOVA produced F(2,149) = 0.73, p =.482, η^2 ≈.010 (trivial). There are no statistically significant differences across IT-knowledge levels; the null hypothesis is retained, with no practically meaningful effect.

Synthesis. The outcomes summarized in Table 7 are consistent with all three hypotheses: no statistically significant differences in the perceived effect of AI tools on the quality of English-language research as a function of gender, age, or IT knowledge. The uniformly very small effect sizes further suggest an absence of practically consequential variation on this dimension within the present sample.

3.3 Conclusion

This study stated that most participants believe AI tools as supported tool in academic research. They agreed that these tools minimize time, ensure quality, and organize the process. At the same time, they also observed some challenges, such as lack of originality and the need for human effort to direct the work.

The results also clarified that the dimensions such as gender, age, IT knowledge, and AI competence did not result in marked differences. While those with advance knowledge

or competence gave slightly higher scores, the overall pattern was the same. Most groups shared the same perspectives and had optimistic attitudes toward using AI in research.

overall, the findings indicate that AI tools are broadly accepted by students and researchers, despite of their personal background. Participants appreciated the enhancement these tools supply, but they also emphasized the significant of balancing AI with human participant.

Chapter Four

Discussions and Conclusions

4.1 Results Related to the Main Question of the Study

What is the effect of AI on the English language research quality as perceived by Palestinian graduate students?

4.2 Results Related to the Main Question of the Study

The evidence shows that AI achieved benefits across the study's three domains. These benefits do not cause any differences between groups related students awareness about quality. Descriptive summaries revealed High central tendency and limited dispersion. Reliability indicators are adequate to support the use of combined scores. Within measurement, the inferential profile is obvious on all domains. Regarding quality/efficiency domain, differences among sub groups were consistently non-significant with small effect sizes. Similarly the strengths/weaknesses domain, the pattern is relatively uniform. Only the strategies domain show significant differences by technical background.

In term of the quality/efficiency domain, the results do not show differential perceptions by gender, age, or IT knowledge. By gender, the independent-samples *t* test showed no significant differences, $t(169) = 0.10$, $p = .92$ (Cohen's $d \approx 0.02$), the effect size is so small that doesn't change anything. By age, a one-way *ANOVA* likewise did not detect differences, $F(2,149) = 1.45$, $p = 0.238$ ($\eta^2 \approx 0.019$). By IT knowledge, the effect was similarly absent, $F(2,149) = 0.73$, $p = 0.482$ ($\eta^2 \approx 0.010$). Considering the high descriptive means, this pattern indicates that basic AI abilities such as automated review, formatting, and citation management are reachable to provide benefit to most students in the same ways.This is actually has role to produce convergent perceptual about quality. To shed light on the situation of graduates in Palestine. Where academic research resources. The effort put in by students is significant and their writing does not match quality standards due to lack of close supervision and access the writing centers.

To turn to the strengths/weaknesses domain, where similarly shows non-significant principal contrasts. For instance, the analysis focusing on age showed $F(2,158) = 1.73$, $p = 0.180$ ($\eta^2 \approx 0.021$), it is again a small effect. The evidence shows that students

alignment in their assessment about AI advantages and disadvantages confirms the perception of AI as a widely accessible support tool that helps solve common writing problems. The results are consistent with the overall pattern: when most scores are high and do not differ much, it becomes harder for statistical tests to detect small differences between groups, so even real effects might not appear.

In contrast, the strategies domain differs by technical background. The evidence shows that students with stronger levels of IT knowledge and AI competence tend to support mature practices at higher levels: $F(2,156) = 6.77$, $p = 0.002$, $\eta^2 \approx 0.080$ for IT knowledge, and $F(2,155) = 5.47$, $p = 0.005$, $\eta^2 \approx .066$ for AI competence. These effects fall within the small-to-moderate range and are meaningful in applied terms, indicating that students with deeper technical literacy are more likely to plan prompts deliberately, cross-check outputs, and document AI-assisted steps transparently.

The literature reviewed earlier provides a balanced external frame for this interpretation. Studies reporting efficiency and clarity gains with guided AI use (e.g., Chien, 2023; Usdan, Chang, & Pensky, 2024) align with the current evidence, suggesting that participants experience tangible and subtle improvements that simplify the writing process. By contrast, research documenting decreased performance under over-dependence or full delegation (e.g., Ju, 2023; Wecks, Voshaar, & Plate, 2024) reduces overgeneralization and suggests a threshold dynamic: strategic use of AI tools in a controlled way can support weaker students in completing tasks, while relying on them too much causes students to depend on them, which leads to reduced learning. In light of this, the benefit derived from artificial intelligence tools depends on the strategy used (how the tools are used by students) rather than on quality standards.

A critical synthesis across the three domains can be summarized as follows. First, the absence of significant subgroup differences in perceived quality does not indicate a lack of relevance; rather, it provides meaningful insight when paired with high means and trivial effect sizes, as it implies non-exclusive access to AI's basic benefits and highlights the limits of global perceptions as a near-term discriminator. Second, the stable analysis of strengths and weaknesses shows that students share a similar awareness of AI's trade-offs. Students recognize AI's support in language while remaining cautious about over-dependence. Third, strategic practices differ by technical background. In this case,

technical background represents a proximal locus, as it is the factor shaping variation among students. What separates students is not whether they use AI, but how they use it. It is worth emphasizing that differential impact primarily appears in students' strategic practices, rather than in immediate differences between groups in perceived overall quality.

The interpretive scope is restricted by significant limitations. The study relies on self-reported data, which may introduce response bias, especially if participants are tech-savvy. Furthermore, the sample size consists of 171 graduate students from West Bank universities, and the cross-sectional design limits the ability to identify causation and leaves an open question: Does the use of AI tools lead to higher research quality, or does higher research quality draw on the use of such tools? Also, providing strong explanations for influence is difficult because of the presence of unmeasured variability, such as tool type, duration of use, and task requirements. In addition, the combination of high means and low dispersion in the descriptive layer suggests possible ceiling constraints that reduce inferential sensitivity to small subgroup effects on global quality.

Within these restrictions, the evidence enhances an accurate conclusion that Palestinian graduate students view AI as a useful academic scaffold. They view it as a tool to decrease burdens and facilitate work procedures.

Moreover, the results in Chapter Three indicate the following:

Domain One: The Effect of AI on Research Quality and Efficiency

The evidence indicates that AI is perceived to provide benefits in research quality and efficiency; yet, those gains do not cause between-group differences in this domain. Descriptive summaries show high central tendency on key indicators of linguistic accuracy, clarity, organization, and workflow, with relatively restricted dispersion; reliability estimates are adequate to warrant interpretation of the composite. Within the measurement context, the inferential profile is consistent. Differences in perceptions across subgroups regarding the quality and efficiency domain are not statistically significant, and the effects are very small. This means that students' evaluations of AI are similar, with no significant differences or multiple levels of evaluation.

The independent samples test shows no significant differences across gender, $t(169) = 0.10$, $p = .92$ (Cohen's $d \approx .02$), reflecting an effect of very small size with no practical value. By age, a one-way analysis similarly does not detect differences, $F(2,149) = 1.45$, $p = 0.238$ ($\eta^2 \approx 0.019$). By IT knowledge, the effect is likewise absent, $F(2,149) = 0.73$, $p = 0.482$ ($\eta^2 \approx 0.010$). For completeness, the analysis by AI competence on the same domain is also non-significant, $F(2,148) = 1.42$, $p = 0.246$ ($\eta^2 \approx 0.019$). The high descriptive means indicate that the basic advantages of AI, such as automated proofreading, formatting consistency, reference handling, and rapid feedback, are apparently reachable and benefit all students in the same way.

A theoretically grounded mechanism helps to make sense of this pattern. AI reduces lower-level cognitive and temporal burdens and thereby frees attention for higher-order tasks; as a result, students report smoother drafting routines and fewer language-level bottlenecks. In a Palestinian graduate context characterized by limited access to writing centers, professional editing, and intensive supervisory feedback, such low-threshold support functions as a compensatory scaffold that brings students to a workable baseline more quickly. When many participants converge on that baseline, global ratings of quality/efficiency tend to cluster near the upper range, and between-group contrasts shrink. The high means and modest dispersion in the descriptive layer are consistent with this interpretation and also imply that conventional tests have reduced sensitivity to small subgroup effects on this outcome when variability is constrained.

Crucially, the absence of significant contrasts on quality/efficiency is not a null result in the pejorative sense. It is substantively informative because it shows broad, non-exclusive access to AI's basic gains and clarifies where differential impact is—and is not—likely to appear in the short term. Elsewhere in the analysis (beyond this domain),students with stronger technical backgrounds implement more mature AI practices, while global quality judgments remain stable across groups. That variation suggests a phased process : AI first reorganizes how work is done (planning, prompting, verification, documentation), then begin the overall evaluation of quality and efficiency, which will appear over a long period and may require time and continuous practice with clear feedback to become measurable with standard instruments.

The broader literature provides a balanced external check on this reading. Findings that guided AI use improves clarity and streamlines writing align with the high perceived gains reported here, and they support a supportive rather than substitutive role for AI in academic production. At the same time, studies documenting diminished performance under heavier reliance or full delegation caution that the benefits are conditional on strategic use. Taken together, these strands are compatible with the present domain-level pattern: students widely recognize AI's contribution to surface quality and workflow, but that recognition remains broadly shared across groups and therefore does not register as significant between-group differences on perceived quality/efficiency.

The interpretive scope of this domain is bounded by several considerations that affect the strength and reach of inference. Outcomes are self-reported, which invites common-method variance and social-desirability inflation; a sampling frame restricted to Palestinian universities with online recruitment limits representativeness and introduces the possibility of self-selection by technology-positive respondents; and the cross-sectional design cannot adjudicate whether higher perceived quality leads to greater AI use or the reverse. In addition, unmeasured heterogeneity in tool type, dosage, instructional guidance, and task demands may attenuate or amplify effects in ways not captured here. Finally, the combination of high *M* and modest *SD* in the descriptive layer points to possible ceiling constraints that suppress detectable subgroup contrasts on global quality/efficiency. Within these bounds, the domain-specific conclusion is precise: AI is widely perceived to enhance research quality and efficiency by alleviating language-level and mechanical burdens, and this enhancement appears broadly shared rather than group-differentiating—an outcome that is coherent with both the local institutional context and the conditional, strategy-dependent benefits identified in recent scholarship.

Supporting studies assist this direction. Bai & Li (2024) discovered that AI writing tools developed the coherence and quality of postgraduate theses. On the other hand, the results should be evaluated with carefulness. Ju (2023) indicates that heavy reliance on AI can decrease accuracy in academic work, as mistakes may go undetected without human review.

Domain Two: Identifying Strengths and Weaknesses of AI Tools

The evidence in this domain indicates perceptual consistency among respondents about what AI 's benefits and limitations lie, but this unanimity does not vary systematically across demographic or technical subgroups. Descriptively, evaluations of AI's advantages (such as grammatical enhancement, stylistic consistency, reference dealing, and workflow Simplifying) occurring together with awareness of drawbacks (such as over- dependence, Superficial reasoning, citation weakness). This balanced pattern is reflected in tightly clustered around high means, suggesting that students shared same understanding rather than division into distinct groups. Inferential tests corroborated this convergence.

The inferential analysis showed that the result is convergent. By gender, the independent-samples result was non-significant, $t(169) = 0.15$, $p = 0.88$, with a trivial magnitude (Cohen's $d \approx 0.03$). By age, a one-way analysis likewise did not detect differences, $F(2,158) = 1.73$, $p = 00.180$ ($\eta^2 \approx 0.021$). Analyses by IT knowledge also yielded non-significant contrasts, $F(2,158) = 1.66$, $p = .193$ ($\eta^2 \approx 0.021$), and the pattern was similar for AI competence, $F(2,157) = 0.76$, $p = 0.470$ ($\eta^2 \approx 0.010$). overall, the minimal effect size suggest that students expressed comparable perspective on AI's strengths and weaknesses, regardless of background.

A theoretically coherent explanation is that students are engaging to a shared set of strengths and weaknesses which are clearly apparent in everyday academic work. the strengths side such as error correction, stylistic refinement, formatting, and quick feedback are readily experienced by most users. All of these lead student to make agreement about benefits. Related to the weaknesses side, all students are aware about common risk such as formulaic language, fabricated references, and avoid analytical effort are equally obvious, These issues encourage a balanced approach that recognizes the usefulness of AI tools while remaining cautious about over dependence. When both sides of the scale are familiar, global judgments of strengths/weaknesses tend to cluster around similar values and differences among sub -groups diminish.

This convergence is also consistent with the larger evidentiary profile of the study. Reliability estimates of data enhances interpreting the composite measure for this domain, and the descriptive layer's high central tendency with modest spread helps

explain why small differences between groups are difficult to detect statistically. Significantly, the flatness here should not be interpreted as analytical vacuity. But, it indicates that short-term differences are unlikely to appear in general evaluations of strength and weakness of AI and in how students apply safeguards against known risks. Indeed, elsewhere in the results, students with stronger technical backgrounds more strongly implement level practices strategy such as Intentional prompting, Clear documentation, validation. Also they keeping themselves aligned with their peers in regard strengths and weaknesses. This disengage has meaningful concept: the society appears to accept on what AI can and cannot do, while Divergent on how to handle its use responsibly.

Compared with prior studies that emphasis this review, the pattern in this domain is obviously balanced. Reports refer that students use AI's tools to support clarity and efficiency align with widely acknowledged advantages. At the same time, results informed that over dependence on AI tools may weaken performance or accuracy. that is help to explain why students also define notable weaknesses. Accordingly, the data support an accurate conclusion in this domain: students express a shared and stable assessment of AI's strengths and weaknesses—a broadly consistent evaluative pattern across gender, age, and technical background—and the lever of differentiation doesn't lie in what they believe but in the strategic practices they adopt to maximize benefits and reduce risks.

From a theoretical standpoint, these results enhance the framework that considers AI as a supportive but non-autonomous partner in academic writing. The awareness of both strengths and weaknesses indicates that students understand the principle that AI can reinforce surface-level writing quality but cannot replace human judgment, innovation, or critical involvement. This perspective Correspond with Chen and Li (2023), who Consider that AI should be Formulated as a collaborative tool rather than an independent agent of authorship.

Nevertheless, interpretation must take into account the study's limitations. The reliance on self-reported perceptions means that students' evaluations of strengths and weaknesses are subjective and may be affected by their own experiences with specific tools. In addition, the Palestinian situation is characterized by irregular internet access

and limited exposure to premium AI applications that enhance concerns regarding accessing reliability. Because the study is cross-sectional, it cannot know how students' perceptions of strength and weakness change over time as they obtain more experience and training.

Overall, the results of the second domain confirmed that Palestinian graduate students perceive AI as a tool with pros and cons.It delivers Significant enhancements in accuracy, clarity, and organization, but it also carries risks of error, over dependence, and Context neglect. This balanced view confirms previous research, extends theoretical understandings of AI as a supportive partner, and underscores the need for training that enable students to maximize AI's strengths while reduce its weaknesses.

Domain Three: Strategies to Maximize the Benefits of AI Tools

Analyses revealed significant differentiation between-group on the strategies domain in related to technical background.. Students with stronger IT knowledge reported higher endorsement of mature AI-use practices, $F(2,156) = 6.77$, $p = 0.002$, $\eta^2 = 0.080$, and students with higher AI competence showed a similar advantage, $F(2,155) = 5.47$, $p = 0.005$, $\eta^2 = 0.066$. These effects fall in the small-to-moderate range and allow for interpretation go beyond mere statistical significance. Unlike the pattered observed for quality /efficiency and on strength/weakness. The current pattern identifies an operational point of differentiation: technical literacy classify students not by whether they use AI, but rather **how** they employ these tools such as purposefully formulating prompts, limiting model scope to task requirements, cross check outputs, ensuring accurate references, documenting AI-assisted steps, and editing with explicit criteria for efficiency.

The configuration of results suggests a staged mechanism. AI's reachable capabilities such as error correction, stylistic refinement, formatting, and quick feedback are largely available and quickly absorbed, which explains consistent perceptions in regard global quality. the discipline of strategic practice separates students, those with stronger technical backgrounds transform general capabilities into optimized workflows that minimize fabrications, control style and form, and fit outputs with the assignment demands. In the Palestinian graduate context, Strategic proficiency operates as a supportive, compensatory framework that directs students' use of the AI tool toward

higher-level objectives rather than limiting them to superficial fixes. The significant effects and their magnitudes enhance this interpretation, while the marginal trends on overall scores elsewhere(Such as total remain close to conventional thresholds) do not contradict it. Global ratings typically delay the development of procedure.

This interpretation is convergent with a balanced strand of prior scholarship. Studies that indicates that students show improvements in clarity, organization, and efficiency, particularly when they are trained to constrain and verify model outputs. The previous research findings corresponds with the current results, which differentiate students based on their strategy adoption (e.g., Chien, 2023; Usdan, Chang, & Pensky, 2024). At the same time, work performance recording that reductions emerged when students rely heavily on or fully delegate tasks highlight that why strategic discipline is determines success rather than volume of use. Experimental evidence viewing losses in accuracy under uncritical reliance on automation. In addition to findings that over-dependence can lead to lower outcomes for high achievers, this is align with the current reasoning that procedural knowledge is the axis on which benefits and risks relay on. (e.g., Ju, 2023; Wecks, Voshaar, & Plate, 2024). Both the internal data and external research enhance a single conclusion: strategy is effective because it sets up operational safeguards such as source checking, prompt formulation aligned with genre standards, and clear documentation which transform AI tools from mere facilitating aids into dependable instruments for academic work.

Evidence from the specific domain also explains why differences between subgroups do not emerge in general evaluations, despite overall positive perceptions. The adoption strategy works as an proximal mediator linking access to AI tools with perceived outcomes. Students with higher knowledge or AI competence are stronger engaging practices that consistently reduces error propagation and stylistic inconsistencies, thereby supporting procedural predictability. Conversely, students with lower technical expertise report fewer such practices and tend to experience AI tools as general aids rather than as controlled instruments. When most students report only surface-level improvements, overall ratings tend to align. however, variations in procedural control remain sufficient to produce the observed effects in the strategies composite. This distinction between immediate procedural actions and broader evaluations accounts for both the significant findings in this study and the stability noted in other contexts.

Finally, The statistical results provide conservative basis for conclusions in this domain. Significant F values with non-trivial η^2 estimates effects indicate that technical expertise is linked to greater support for practices that enhance the advantages of AI's tools while reducing potential drawbacks. The data reinforce the view that strategic discipline is what distinguishes students in their level of AI engagement. This disciplined approach becomes the key differentiating factor among students especially Within a context where academic students encounter structural limitations on their abilities to engage with AI tools. In context where academic writing support is limited, this procedural pathway appears theoretically coherent and supported by evidence. Under conditions of scarce resources, the effectiveness of students strategic choices becomes the decisive factor that determines whether AI tool fosters deep academic development or merely as a quick convenience that plateaus only superficial gains.

General Conclusion on the Three Domains

Taken together, the three-domain profile is internally coherent and theoretically disciplined. Perceived **quality/efficiency** and the inventory of **strengths/weaknesses** display consistently high central tendency with modest dispersion, and inferential tests on these two domains were uniformly non-significant with trivial effect sizes (e.g., $t(169) = 0.10$, $p = 0.92$, Cohen's $d \approx 0.02$; $F(2,149) = 1.45$, $p = 0.238$, $\eta^2 \approx 0.019$). The data therefore support convergence rather than stratification in how students appraise AI's contribution to English-language research quality and in how they balance its advantages and liabilities. This convergence is substantively meaningful: low-threshold affordances—error correction, stylistic polishing, formatting, and rapid feedback—are broadly accessible and quickly internalized, particularly in a resource-constrained graduate environment, which raises many students to a similar functional baseline and compresses between-group variation in global judgments.

By contrast, the strategies domain identifies a procedural locus of differentiation aligned with technical background. Students reporting stronger IT knowledge and higher AI competence endorsed more mature practices—deliberate prompting, source-based verification, citation hygiene, and transparent documentation—with statistically significant differences and small-to-moderate effect sizes (e.g., $F(2,156) = 6.77$, $p = 0.002$, $\eta^2 \approx 0.080$; $F(2,155) = 5.47$, $p = .005$, $\eta^2 \approx 0.066$). The configuration across domains thus supports a staged mechanism: AI first reshapes how students execute writing tasks

(procedural literacy) while global evaluations of quality and of pros/cons remain broadly shared. As procedural control consolidates—through better prompting and systematic verification—downstream differences in objectively rated quality could emerge, but the present evidence indicates that such divergence is not yet visible in perceived, global outcomes.

This pattern is consistent with a balanced reading of prior scholarship. Studies that report improvements in clarity, organization, and efficiency under guided use align with the high means and the widespread recognition of advantages, whereas findings of diminished performance and accuracy under heavy reliance or uncritical delegation explain why respondents also register salient weaknesses. The present data reconcile these strands by locating variance at the strategy level rather than in general attitudes: students largely agree about what AI can and cannot do, yet differ—modestly but meaningfully—in the discipline with which they control and audit its outputs.

Critically, the absence of significant subgroup contrasts on quality/efficiency and on strengths/weaknesses is not an analytical void. High descriptive means and minimal effect sizes d/η^2 indicate that AI basic benefits extend across all groups, producing largely uniform perceptions in global evaluations. At the same time, the significant differentiation in strategies points to the area where specific improvements are occurring now and where future divergence is most likely to develop, specifically in the establishment of systematic, fine-tuned workflows that reduce hallucinations, manage register and genre, and ensure alignment with disciplinary standards.

The interpretive scope remains constrained by well-known threats to valid inference including self-reported measures that are susceptible to common-method bias, an online graduate cohort with restricted generalizability, a cross-sectional design, and unobserved variation in tool type, intensity of use, level of guidance, and task requirements. Descriptive that cluster near the upper range also suggest potential ceiling constraints, which reduce sensitivity to small subgroup effects on global outcomes. Within these bounds, the general conclusion is precise: AI is widely perceived to enhance research quality and efficiency and to offer a recognizable set of strengths (alongside acknowledged risks), and differential impact resides primarily in strategic practice rather than in immediate, between-group divergence on overall quality. This domain-level

synthesis explains both the breadth of perceived benefits and the locus of emerging differentiation, providing a coherent foundation for the study's subsequent discussion and implications.

4.3 Discussion of the results of the hypothesis

4.3.1 Results related to the first hypothesis

There are no statistically significant differences in the perceived effect of AI tools on the quality of English language research among Palestinian graduate students according to the ender variable.

The first hypothesis posits that the perceived effect of AI on English-language research does not differ by gender. The inferential profile supports this statement across all study domains. On the quality/efficiency axis, the independent-samples test was non-significant, $t(169) = 0.10$, $p = 0.92$ (Cohen's $d \approx 0.02$). A similarly flat pattern holds for the strengths/weaknesses axis, $t(169) = 0.15$, $p = 0.88$ ($d \approx 0.03$), for the strategies axis, $t(169) = 0.90$, $p = 0.37$ ($d \approx 0.17$), and for the total score, $t(169) = 0.75$, $p = 0.45$ ($d \approx 0.15$). Effect sizes are uniformly trivial to small. Read together, these statistics indicate convergence rather than stratification in how male and female graduate students appraise AI's contribution to research quality, its advantages and limitations, and the practices they endorse when using it.

Interpretively, the pattern is coherent with the functional character of the tools under study. Core affordances—automated proofreading, stylistic polishing, formatting and reference handling, and rapid feedback—are low-threshold and broadly accessible. When such affordances are widely available, students of different genders arrive at a similar operational baseline: they draft faster, resolve language-level bottlenecks more efficiently, and experience fewer mechanical disruptions to argument development. Under these conditions, global judgments of quality and efficiency cluster near the upper range, and between-group contrasts shrink. The descriptive layer of the data (high central tendency with modest dispersion) enhance this reading and clarifies why inferential sensitivity to small subgroup effects is limited when variability is restricted.

The absence of gender differences on the strengths/weaknesses axis is equally informative. It appears that respondents tend to agree on a detailed framework of

development: they are aware of the tangible improvements in language skills and workflow, while also recognizing risk associated with it such as phrasing, reference fragility, and the temptation to offload analytic effort. This shared understanding indicates how both strengths and weakness are clear in daily practice academic work and explains the compressed dispersion around high means. In short, students do not divided into gendered groups about if AI is good or bad. They express essentially the same set of trade-offs.

The strategies axis follows same principle. Even though notable differences appear according to students technical expertise in the study, gender does not determine influence on how students implement AI use practices including source-driven verification, prompts crafted according to genre conventions, and transparent record-keeping. The non-significant result, $t(169) = 0.90, p = 0.37$ ($d \approx 0.17$), indicates alignment rather than divergence in procedural discipline across genders. This matters because strategy is usually what causes differences. the absence of such an effect here support the conclusion that gender does not serve as the mechanism influencing the seriously students' engage with AI tools.

A critical reading also considers test power and measurement quality. With $n \approx 171$ and balanced groups, the design is positioned to detect medium effects; The consistently very small d values indicate that there are no significant differences related to gender that they may have missed. Reliability estimates for the composite measures are sufficient so it is reasonable comparisons credibility at the level. at the same time, high means with restricted spread descriptive layer refer to suggest ceiling effects that inherently reduce the observable difference in global indices. While this statistical reality constrains the generalizability of the findings, it does not diminish the significance conclusion of gender alignment.

Contextual factors reinforce the interpretation. In Palestinian graduate settings, access to writing centers, professional editing, and intensive supervisory feedback is constrained. AI thus functions as a compensatory scaffold that delivers similar low-threshold gains to a broad swath of students, irrespective of gender. When a technology equalizes bottlenecks that are common to all—lexical accuracy, consistency, and local

coherence—perceptions of quality and of the tools' pros and cons are expected to homogenize across groups that differ on other dimensions.

The inferential conclusion is therefore precise: the evidence shows no statistically significant gender differences in perceived AI impact on quality/efficiency, on the articulation of strengths and weaknesses, or on the endorsement of strategy-level practices, with effect sizes that are trivial to small across outcomes. This is a substantively positive result in the sense that it indicates non-exclusive access to AI basic benefits and a shared, critically aware stance toward its liabilities. It is also a methodologically disciplined result: the statistics reported above, interpreted against the descriptive and reliability backdrop, justify retaining the null hypothesis for gender without equivocation.

The scope of this conclusion remains bounded by familiar threats to inference that are addressed in the study as a whole. Outcomes are self-reported, which introduces common-method variance and potential social-desirability inflation; the online graduate sample limits representativeness and invites self-selection toward technology-positive respondents; the cross-sectional design cannot adjudicate directionality between perceived efficacy and AI use; and unmeasured heterogeneity in tool type, dosage, guidance, and task demands complicates strong mechanistic claims. Even within these constraints, the pattern for the first hypothesis is consistent and theoretically credible: gender does not structure how students evaluate AI contribution, because the prominent benefits are low-threshold and shared, while meaningful differentiation—where it exists elsewhere in the study—tracks technical literacy rather than gender.

4.3.2 Results related to the second hypothesis

The second hypothesis states that the perceived effect of AI on English-language research does not differ across age groups. The inferential results support this claim across all study domains. On the quality/efficiency axis, a one-way analysis of variance was non-significant, $F(2,149) = 1.45$, $p = 0.238$, $\eta^2 = 0.019$. A similarly flat pattern emerged for the strengths/weaknesses axis, $F(2,158) = 1.73$, $p = 0.180$, $\eta^2 = 0.021$, and for the strategies axis, $F(2,156) = 1.03$, $p = 0.361$, $\eta^2 = 0.013$. The total composite likewise showed no age effect, $F(2,166) = 1.19$, $p = 0.306$, $\eta^2 = 0.014$. In each case, effect sizes are very small, indicating that any between-group differences—if present—are trivial in practical terms.

Interpreted against the descriptive and reliability backdrop established earlier, this profile indicates convergence rather than stratification in how graduate students of different ages appraise AI contribution to research. High central tendency with modest dispersion on the descriptive layer suggests that widely accessible affordances—automated proofreading, stylistic polishing, formatting consistency, reference handling, and rapid feedback—are experienced in broadly similar ways across cohorts. When a technology equalizes common bottlenecks in the writing workflow, global ratings of quality and efficiency tend to cluster near the upper range, and inferential contrasts by age naturally shrink. The adequacy of internal consistency estimates for the composites further supports treating these non-significant findings as substantively informative rather than as artifacts of noisy measurement.

A theoretically grounded mechanism clarifies why age fails to differentiate outcomes while technical background does so elsewhere in the study. The benefits most salient to respondents are low-threshold and tool-driven; they do not require age-specific experience so much as minimal familiarity with interface conventions that are now pervasive in academic life. In graduate programs, the functional literacy needed to engage basic AI affordances is widely distributed, and the typical age span is comparatively narrow. Both conditions pull perceptions toward a common operational baseline. In contrast, procedural proficiency—deliberate prompting, scope control, verification against sources, and transparent documentation—is not age-dependent per se but strategy-dependent, which explains why age does not sort students on the strategies axis while IT knowledge and AI competence do. The present results therefore disentangle *who students are* (by age) from *how they work* (by strategy), and only the latter emerges as a locus of differentiation in the broader analysis.

From a methodological perspective, The statistical its clearly show support for keeping the null hypothesis. With the three age categories (as indicated by the degrees of freedom) and the observed variance structure, the study design is sufficiently able to detect effects of medium size. The consistently very small η^2 values show that there are no meaningful differences related to age that that would have been missed by underpowered testing. Simultaneously, the descriptive pattern, characterized by high means and limited variability, indicates to reduce sensitivity into small differences in global indices.

Contextual factors in the Palestinian graduate environment reinforce this interpretation. Limited access to writing centers, professional editing, and intensive supervisory feedback means that AI often functions as a compensatory scaffold that streamlines lower-level tasks for all users, irrespective of age. When the key gains involve reducing cognitive and temporal load and smoothing recurrent language-level bottlenecks, age cohorts are expected to converge in their evaluations. What differentiates users, where it does so elsewhere in the data, is the discipline of strategic practice, not demographic position in the life course.

The interpretive scope remains bounded by the same threats to inference addressed for the first hypothesis. Outcomes are self-reported, inviting common-method variance and socially desirable responding; an online, program-based sampling frame limits representativeness and introduces the possibility of self-selection; the cross-sectional design cannot adjudicate directionality between perceived efficacy and AI use; and unmeasured heterogeneity in tool type, dosage, guidance, and task demands complicates strong mechanistic claims. Even within these constraints, the evidence is internally consistent and theoretically credible: **age does not structure** how Palestinian graduate students evaluate AI's impact on research quality, on the perceived balance of strengths and weaknesses, or on the endorsement of strategy-level practices. The pattern fits a coherent account in which AI provides low-threshold support that is broadly shared across cohorts, while meaningful differentiation—where it appears in the study—tracks technical literacy and strategic control, not age.

4.3.3 Results related to the third hypothesis

There are no statistically significant differences in the perceived effect of AI tools on the quality of English-language research among Palestinian graduate students according to the IT-knowledge variable.

The third hypothesis asserts that the perceived effect of AI tools on the quality of English-language research does not differ by IT-knowledge level. The inferential evidence supports this claim on the quality/efficiency axis. A one-way analysis of variance was non-significant, $F(2,149) = 0.73$, $p = .482$, $\eta^2 = .010$, indicating a trivial effect size that lacks practical import. The broader profile is consistent with this pattern: the strengths/weaknesses axis likewise yielded a non-significant contrast, $F(2,158) = 1.66$, p

=.193, η^2 =.021, while the total composite showed only a trend that does not meet the decision threshold, $F(2,166) = 2.72$, p =.069, η^2 =.032. Considering high descriptive means and modest dispersion, the findings indicate that there is no statistically significant effect of IT knowledge on perceived respect to research quality, thus supporting the retention of the null hypothesis.

The configuration of results is theoretically coherent. Basic AI capabilities such as automated review, stylistic polishing, formatting, and reference handling are enough reachable to support students in same way. With these abilities students experience show similar improvements in drafting speed, linguistic accuracy, and workflow stability. Consequently, overall evaluation judgments of quality tend to cluster at higher levels, reducing variability between groups and limiting the potential for detecting differences across IT-knowledge tiers. Reliability estimates for the composites warrant confidence that these non-significant findings are not artifacts of unstable measurement; rather, they indicate substantive homogeneity in how students at different IT-knowledge levels appraise AI contribution to quality.

Considering comprehensive analysis determines where differentiation does occur: on the strategic axis, students with stronger AI experience tend to adopt more advanced practices, careful prompting, checking sources, proper citations, and clear documentation, with statistically significant yet modest effects. Basically, AI shape the way students work first, whereas overall quality judgments remain largely consistent across IT knowledge. This sequencing explains why the hypothesis is supported regarding perceived quality even as strategy adoption differentiates respondents in other aspects in the study.

A critical reading of the measurement context sharpens the conclusion. Descriptive summaries show high central tendency with relatively restricted spread, a structure that implies ceiling constraints on global indices and reduces sensitivity to small subgroup effects. Within that structure, the very small η^2 values reported above argue against hidden medium effects that the design failed to detect. Moreover, the outcomes in this section are self-reported perceptions of quality, not blind, rubric-based ratings of manuscripts. The evidence therefore supports a precise claim: students across IT-knowledge tiers converge in their *perceptions* of AI's contribution to quality, while the locus of

differential impact lies in procedural discipline, not in immediate divergence on global quality judgments.

Contextual features of Palestinian graduate study reinforce this interpretation. Where access to writing centers, professional editing, and intensive supervisory feedback is constrained, AI functions as a compensatory scaffold that levels common bottlenecks—lexical accuracy, consistency, and local coherence—across the IT-knowledge distribution. When a technology equalizes such bottlenecks, perceived quality tends to homogenize. The observed pattern is thus substantively positive: non-exclusive access to AI's basic benefits produces shared improvements in the writing workflow without stratifying students by IT-knowledge level on the quality/efficiency axis.

Within these bounds, the third hypothesis is adjudicated without equivocation: the data support retaining the null for perceived research quality by IT knowledge ($F(2,149) = 0.73$, $p = .482$, $\eta^2 = .010$), with corroborating non-significant findings on related domains and only trend-level movement on the total composite. The evidence delineates clear boundary conditions for near-term impact in this population—broadly shared gains in perceived quality, coupled with meaningful differentiation in strategy adoption—and it specifies where subsequent inquiry should look for downstream divergence in objectively rated outcomes.

4.3.4 Results related to the fourth hypothesis

There are no statistically significant differences in the perceived effect of AI tools on the quality of English language research among Palestinian graduate students according to the AI experience variable.

The fourth hypothesis states that the perceived effect of AI tools on the quality of English-language research does not differ across levels of AI experience. The inferential evidence supports this claim on the quality/efficiency axis: a one-way analysis of variance was non-significant, $F(2,148) = 1.42$, $p = 0.246$, $\eta^2 = 0.019$, which indicates a trivial effect size with no practical import. Convergent non-significant patterns appear on the strengths/weaknesses axis, $F(2,157) = 0.76$, $p = 0.470$, $\eta^2 = 0.010$, and the overall composite shows only a trend that does not meet the decision threshold, $F(2,165) = 2.73$, $p = 0.068$, $\eta^2 = 0.032$. Taken together with the high descriptive means and modest

dispersion, these results justify retaining the null hypothesis for perceived research quality by AI experience.

This profile is theoretically consistent with the functional character of the tools. Once students achieve minimal functional literacy, low-threshold affordances—automated proofreading, stylistic polishing, formatting consistency, reference handling, and rapid feedback—are broadly accessible and quickly internalized. Under these conditions, students with different levels of AI experience demonstrated similar surface-level improvements in drafting fluency and workflow stability, thereby reduce between-group variability in overall judgments of quality and efficiency. Reliability of the composite scores is adequate, so the non-significant results are better understood as meaningful Homogeneity rather than measurement noise.

In parallel, the wider analysis reveals a practical difference that is not part of the main hypothesis. Within the strategic axis—where the outcome reflects how students plan, prompt, verify, and document using AI. technical familiarity plays a differentiating role: respondents with stronger technical literacy are more likely to adopt advanced methodological practices, $F(2,155) = 5.47$, $p = 0.005$, $\eta^2 = 0.066$. This difference isn't a contradiction since it actually helps explain what's happening. The results show that AI experience improves how students work, without producing differences in overall quality perceptions. Usually, Procedural improvements tend to occur first and then later changes emerge in global quality evaluations, especially when those overall ratings are already concentrated at the higher end of the descriptive scale.

Context is crucial for understanding this pattern. Graduate students in Palestine have limited support for writing and editing, the AI tool fills that gap and helps them deal with language difficulties, regardless of their experience level. When a technology equalizes common obstacles, perceived quality converges. What differs, where it does, is the discipline of control: experienced users more consistently constrain model scope, verify outputs against sources, maintain citation hygiene, and document AI-assisted steps, while less experienced users benefit primarily from generic surface support.

Methodologically, the statistics reported above warrant a precise adjudication. With three experience levels (as reflected in the degrees of freedom) and the observed variance structure, the design is positioned to detect medium effects; the very small η^2 values argue

against latent, practically meaningful differences on perceived quality that would have emerged with greater power. Descriptive clustering near the upper range suggests ceiling constraints that naturally reduce sensitivity to very small between-group effects on global indices; that constraint defines the scope of inference without undermining the substantive conclusion.

In sum, the evidence shows no statistically significant differences in the perceived effect of AI tools on English-language research quality across AI-experience levels ($F(2,148) = 1.42$, $p = 0.246$, $\eta^2 = 0.019$), with corroborating null findings on related domains and only trend-level movement on the total composite. The impact of experience is primarily appear at the procedural level in through differences in strategy use. rather than its variation to overall quality strategies. This pattern aligns with theoretical expectations and fit credibly within the context of this study.

4.4 General Conclusion of Chapter Four

This chapter combine descriptive, reliability, and inferential evidence to characterize how Palestinian graduate students understand and use AI in English language research across three domains, quality/efficiency, strengths/weaknesses, and strategies, and to adjudicate the study's hypotheses. Overall, students scored high on average with only small differences, and the measures were reliable. While general ratings of quality and efficiency were similar across groups, how students adopted strategies varied noticeably based on their technical background.

Regarding quality/efficiency, there were no significant differences between students based on age, gender, IT knowledge or AI experience. Effect sizes (Cohen's d and η^2) were very small. Considering the high averages and limited variability in the descriptive data, these non-significant results suggest that students generally agree on how AI contribute to the overall quality and fluency of their research writing. This is probably AI capabilities such as automated proofreading, stylistic editing, consistent formatting, reference management, and quick feedback are easily accessible and quickly learned, especially in settings with limited formal writing support. When most students reach the same basic level, overall ratings end up high and similar, which matches the very small effect sizes.

A similar reasoning applies to the strength /weakness domain. Participants provided a shared, nuanced evaluation of AI's strengths and weaknesses. They recognized clear benefits for language use and workflow but also pointed out familiar problems, like repetitive phrases, weak references, and over reliance on AI' tools. The limited variability in responses is meaningful rather than meaningless; it indicates that most people share similar views about the pros and cons of AI, which explains why most scores were high and close together.

The strategic domain identified a specific procedural point where differences emerged and this variation was related with technical background. Students with stronger IT knowledge and AI experience demonstrated more advanced AI use practices, like asking clearer prompts, setting limits, checking sources, and keeping clean citations,with statistically significant differences and small to moderate magnitudes (F and p values met conventional thresholds; η^2 estimates were on the order of.066–.080)the stability observed in studies support a staged mechanism: AI first change how students doing writing tasks (procedural literacy), while global evaluations of quality, weakness and strength remain shared broadly. Where the data hinted at movement in the overall index, the signals remained at trend level and did not reach the decision threshold. That is aligned with the idea that overall judgments of quality tend to improve more slowly than the maturation of procedure.

Meaningful subgroup differences appeared in perceived quality across gender, age, IT knowledge, and AI experience. The conclusion supported by consistent findings in related areas and by very small effect sizes. At the same time, technical background influenced the strategic adaptation, marking the divide between broadly shared, superficial improvements and the more limited development of structured, disciplined practice. The pattern makes sense theoretically and fits well with the surrounding contest in Palestinian graduate sitting where reaching sources is limited, AI tools work as support to help students deal with language challenges. However meaningful differences appear once students start to develop stronger procedural skills.

The conclusions can be interpreted is limited by familiar constraints explained earlier in the chapter. Outcomes were self-reported rather than blind, rubric-based ratings, which introduces common-method variance and possible social-desirability inflation; the online,

program-based sampling frame limits representativeness and invites self-selection; the cross-sectional design precludes causal attribution; and unmeasured heterogeneity in tool type, dosage, guidance, and task demands complicates strong mechanistic claims. Additionally, the combination of high *M* and modest *SD* suggests possible ceiling constraints that reduce sensitivity to small subgroup effects on global indices. Within these bounds, the chapter's general conclusion is precise: students widely perceive AI to enhance research quality and efficiency and to offer a recognizable set of strengths alongside acknowledged risks; differential impact resides primarily in strategic practice, not in immediate, between-group divergence on perceived overall quality. This domain-level synthesis provides a coherent foundation for the study's final implications and the directions articulated in the concluding sections.

4.5 Recommendation

According to the findings across the three domains, the following recommendations are proposed for different stakeholders.

4.5.1 Recommendations for Universities and Colleges

The following recommendations are suggested:

- **Incorporating AI into research training:** Embed a short "Responsible AI in Research Writing" module inside research-methods and academic-writing courses that includes genre-specific prompt design (for literature review, methods, results, discussion), a clear definition of the scope and limitations (what AI may and may not do), writing with *AI disclosure* language, citation hygiene, and verification routines (fact-checking claims, triangulating sources, matching DOIs), assessed through authentic pre/post writing samples.
- **Providing institutional access:** that involves establishing a single sign-on portal for approved digital tools and secure equitable availability through ministry/consortium licensing to reduce per-seat cost, need-based fee waivers/vouchers, extended computer-lab hours and device-lending for students without hardware, low-bandwidth materials on the LMS, and endorsed zero-cost alternatives (e.g., Zotero) with clear privacy controls and data-handling guidance.

- **Developing practical workshops:** Offer sessions series that (1) teaches disciplined drafting and outline-to-prompt techniques; (2) focuses on reviewing for clarity, cohesion, and register control; (3) builds solid reference-management workflows (import, DOI matching, fixing AI-wrong citations); and (4) trains verification and transparency through a simple *Verification Log* and a concise *AI Disclosure Statement*, with take-home templates and brief practice datasets.
- **Promoting critical engagement:** Require a short reflection with each major assignment that recognizes AI-assisted edits from the student's own analysis and explains with evidence the student's own point of view, thereby reinforcing originality, argument quality, and proper source use rather than substituting human reasoning with automated text.
- **Handling contextual constraints:** Publish bilingual quick guides with Palestinian examples, arrange evening/weekend low-cost or free clinics with the local language centers and international NGOs, and adopt a train-the-trainer model that equips teaching assistants to deliver discipline-aware support at a wide range.
- **Standardizing pre-submission quality controls:** Introduce a brief checklist (attached to every submission) covering AI disclosure, verifying reference, keeping style consistency, and alignment with assignment objectives, so that improvements in efficiency are matched by definite scholarly practice.
- **Creating a privacy-rated tool list:** introducing a living catalogue of approved tools with summarized data-privacy conditions and intended use cases (editing vs. generation) so students can choose responsibly without incurring additional costs.

4.5.2 Recommendations for Faculty Members

- **Raising critical awareness:** Deliver a required mini-unit using a *failure-modes casebook* that demonstrates common AI pitfalls (inaccurate facts, fabricated references, genre/register drift, and cultural insensitivity) and has students diagnose and correct them using documented steps.
- **Developing clear ethical guidelines:** Issue a one-page policy that classifies uses as allowed, allowed with disclosure, or prohibited; mandates a concise AI-use statement in submissions; and commits to human, evidence-based review of integrity cases (with AI detectors treated as advisory, not determinative).

- **Promoting balanced integration:** Redesign key assignments so AI can assist surface-level language tasks while the graded core remains human analysis (interpretation of findings, theoretical integration, critique), evidenced by the verification log rather than generic, automated summaries.
- **Designing specialized workshops:** Provide discipline-specific labs (e.g., TESOL, linguistics, education) that contrast appropriate vs. inappropriate AI outputs, train students to repair citations and claims, and practice controlling tone and genre so outputs fit disciplinary expectations.
- **Supporting contextual adaptation:** Publish a bilingual style addendum that localizes terminology, institutional names, and policy references, and demonstrates how to adapt AI-generated text to Palestinian research contexts and field-specific conventions.
- **Institutionalizing transparent practice:** Add a short honor pledge inside the LMS submission flow that acknowledges the policy, confirms disclosure, and links to the verification log, thereby normalizing responsible use without adding financial burden.
- **Monitoring understanding over time:** Run brief pulse surveys and periodic writing audits to track growth in students' critical awareness of AI's strengths and weaknesses, feeding the results back into workshops and course design.

4.5.3. Recommendations for the Ministry of Education

- **Providing structured strategy training:** Publish a simple Tool–Task Matrix that maps research stages (planning, drafting, revising) to approved tools and requires a *Verification Log* with each submission, ensuring students apply the right tool at the right step and can demonstrate control over outputs.
- **Encouraging collaborative practices:** Build rubric-based peer-review cycles and scheduled supervisor checkpoints that prioritize argument quality, evidence integrity, and disciplinary fit, so that AI assistance is continuously triangulated with human feedback.
- **Promoting verification practices:** Offer library-led clinics that teach database search, DOI/ISBN matching, and source triangulation; require exporting reference

lists from Zotero/Mendeley rather than pasting AI-generated citations; and set clear expectations that every factual claim is checkable.

- **Integrating AI literacy into methods courses:** Assess a compact portfolio comprising the initial prompt(s), an excerpt of AI output, the student's human edits and rationale, evidence of verification, and a short AI disclosure paragraph, so strategic discipline is learned and demonstrated—not just stated.
- **Institutionalizing ethical frameworks:** Publish an acceptable-use table tailored to research writing and pair it with a brief pre-submission checklist; together they make the expected strategy (prompting, verification, documentation) explicit and auditable.
- **Building a prompt and pattern library:** Maintain a growing repository of vetted prompts and genre patterns (with "before/after" examples) that students can adapt, reducing trial-and-error costs while reinforcing disciplined workflows.
- **Recognizing disciplined practice:** Offer a small micro-credential or badge for completion of the strategy pathway (workshops + portfolio + verified checklist), signaling readiness to use AI responsibly without imposing additional fees.
- **Financial Access:** Universities can support equitable access by negotiating consortium licenses at the ministry or multi-university level; issuing means-tested waivers/vouchers; extending lab hours and device-lending; endorsing robust, zero-cost tools (especially for reference management); and providing bilingual, low-bandwidth materials on the LMS—thus ensuring that training, workshops, and responsible-use practices are accessible to all students regardless of economic background.

4.5.3 Recommendations for Students

- Adopt **ethical and responsible** AI practices, verifying outputs, and keeping academic integrity.
- Use AI as a **supportive** tool rather than a replacement for critical thinking.
- Engage in **structured strategies** such as disciplined prompting, cross-checking information, and documenting AI-assisted steps.
- Develop **awareness of AI strengths and weaknesses** to avoid over dependence and guarantee meaningful learning.

4.5.5 Recommendations for Future Research

- Conduct **longitudinal studies** to examine the long-term impact of AI use on research quality and academic development.
- Examine how **variables** such as tool type, duration of use, task demands, and discipline influence AI effectiveness.
- Expand **samples to include different regions**, institutions, and academic levels to enhance generalizability.
- Investigate **instructional and policy interventions** that best support ethical, strategic, and effective AI use.

References

Abdulwahab, H., Mahmoud, A., & Rashwan, S. (2023). The Applications of Artificial Intelligence and Their Impact on Developing Linguistic Creativity Among Gifted Secondary School Students. *Journal of Faculty of Education, 39*(1), pp. 109 - 135.

Alhajji, R. (2024). Assessing the influence of AI on modern student writing standards: An educators'. *Research Journal in Advanced Humanities, 5*(3). Retrieved from https://doi.org/10.58256/5tpepv73.

Al-Hakami, A. (2023). Artificial intelligence applications in Saudi public education: Opportunities and challenges. *Journal of Educational Technology and Policy, 15*(2), pp. 112 - 130.

Alharbi, W. (2023). AI in the Foreign Language Classroom: A Pedagogical Overview of Automated Writing Assistance Tools. *Education Research International*, p. 15. Retrieved from https://doi.org/10.1155/2023/4253331.

Almeida, L. (2025). Measuring reliability in educational research: New trends and recommendations. *Journal of Educational Measurement, 26*(1), pp. 45-59. Retrieved from https://doi.org/10.3102/edmeasure2025.

Alnafea, S. (2018). *Robot programme for the talented: Educational robot programming skills for gifted learners through simulation-based software.* Jedda, Saudi Arabia: King Abdulaziz University.

AlQarna, S. (2021). The Effect of Artificial Intelligence (Microbit) in Raising the Motivation towards Learning Programming Among the Students of Educational Technology at King Abdulaziz University in Jeddah. *Journal of Educational and Psychological Sciences, 30*, pp. 58-76.

Alshamy, A., Al-Harthi, A. S., & Abdullah, S. (2025). Perceptions of Generative AI Tools in Higher Education: Insights from Students and Academics at Sultan Qaboos University. *Education Sciences, 15*(4). doi:https://doi.org/10.3390/educsci 15040501.

Badah, A., Najjar, S., Almimi, A., & Alqbailat, N. (2024). Language Challenges Encountered by Palestinian Postgraduate Students in Academic Written Discourse. *l'analisi linguistica e letteraria*, pp. 93-114. Retrieved from https://www.analisilinguisticaeletteraria.eu/index.php/ojs/article/view/557/464.

Bai, L., & Li, X. (2024). The role of AI writing tools in improving postgraduate research quality. *Journal of Academic Writing, 14*(2), pp. 113 - 129.

Chen, Y., & Li, M. (2023). Human–AI collaboration in academic writing: Balancing efficiency and originality. *Journal of Academic Writing Research, 14*(2), pp. 75 - 92.

Chien, S. (2023). The impact of AI writing tools on English academic writing among non-native speakers. *System*.

Coleman, J. (2020). *AI and Our Understanding of Intelligence. Intelligent Systems with Applications.*

Descriptive Research in Psychology. (2024). Verywell Mind. Retrieved from https://www.verywellmind.com/descriptive-research-in-psychology-8551840.

Doronila, R. (2025). Is the Likert Scale Still IN in 2025? Revisiting Usage, Interpretation, and the Rounding Paradigm. doi:10.13140/RG.2.2.16828.76161.

Einarsson, H., Lund, S. H., & Jónsdóttir, H. A. (2024). Application of ChatGPT for automated problem reframing across academic domains. *Computers and Education: Artificial Intelligence, 6*. Retrieved from https://doi.org/10.1016/j .caeai.2023.100194.

Fitria, F. (2021 b). ProWritingAid as AI-powered writing tools: The performance in checking grammar and spelling of students' writing. *ResearchGate*.

Gutiérrez, L. (2023). Artificial Intelligence in Language Education: Navigating the Potential and Challenges of Chatbots and NLP. *Research Studies in English Language Teaching and Learning, 1*(3), pp. 180 - 191. Retrieved from DOI:10.62583/rseltl.v1i3.44.

Hamdan, R., Ashour, W., & Daher, W. (2021). The role of the E-Learning departments in controlling the quality of electronic assessments in Palestinian universities during the COVID-19 Pandemic. *Sustainability, 13*(21). doi:https://doi.org/10.3390/su132112021.

Heydarifard, Z., & Jahanbin, A. (2025). Questionnaire development quality assessment: The curricula and courses in health. *Education Journal, 14*(3), pp. 154 - 159. doi:https://doi.org/10.11648/j.edu.20251403.18.

Huang, X., Zou, D., Cheng, K., & Chen, X. (2023). Trends, research issues and applications of artificial intelligence in language education. *Educational Technology & Society*, pp. 112-131.

Huang, W., Nurtantyana, R., Purba, S., Hariyanti, U., Indrihapsari, Y., & Surjono, H. (2023). AI and Recognition Technologies to Facilitate English as Foreign Language Writing for Supporting Personalization and Contextualization in Authentic Contexts. *Journal of Educational Computing Research*, pp. 1008 - 1035. Retrieved from https://doi.org/10.1177/07356331221137253.

Johnson, R., Lee, A., & Martinez, P. (2024). Online surveys in times of crisis: Opportunities and challenges for educational research. *International Journal of Social Research Methodology, 27*(2), pp. 145 - 160. Retrieved from https://doi.org/10.1080/13645579.2023.2001234.

Ju, Q. (2023). Experimental evidence on negative impact of generative AI on scientific learning. Retrieved from https://doi.org/10.48550/arXiv.2311.05629.

Kim, Y., Lee, M., Kim, D., & Lee, S. (2023). Towards explainable AI writing assistants for non-native English speakers. Retrieved from https://arxiv.org/abs/2304.02625.

Kurniati, E., & Fithriani, R. (2022). Post-Graduate students' perceptions of Quillbot utilization in English Academic Writing class. *Journal of English Language Teaching and Linguistics, 3*. Retrieved from https://doi.org/10.21462/jeltl.v7i3.852.

Liu, H., & Wang, S. (2025). Selective integration of AI tools in postgraduate research: Avoiding dependency and improving organization. *Educational Technology Research and Development, 73*(2), pp. 221 - 239.

Liu, X., & Zhang, Y. (2025). The role of AI-assisted learning in academic writing: A mixed-methods study on international ESL students. *Pedagogies Journal, 27*(2), pp. 115 - 132.

Liu, Y., & Zhang, H. (2025). AI-based grammar checkers and the improvement of academic writing among international students. *Journal of English for Academic Purposes*.

Liu, Y., Chen, L., & Yao, Z. (2022). The application of artificial intelligence assistant to deep learning in teachers' teaching and students' learning processes. *Frontiers in Psychology*, pp. 1-13. Retrieved from https://doi.org/10.3389/fpsyg.2022.929175.

Marzuki, M., Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education, 10*(2). doi:https://doi.org/10.1080/2331186X. 2023.2236469.

Merisi, S., & Mtanha-Matariro, N. (2024). Uses and challenges of AI in academic writing in higher education. *Stanford Journal of Applied Research, 4*(1).

Mohammed, A. (2024). The impact of artificial intelligence applications on enriching English language teaching among primary school students in Aleppo. *Journal of Modern Educational Research, 12*(1), pp. 55 - 72.

Mu, P. (2019). Research on Artificial Intelligence Education and Its Value Orientation. *1st International Education Technology and Research Conference (IETRC 2019).* China. Retrieved from https//webofproceedings.org.

Musa, H., & Almryad, A. (2025). Comparison between five popular paraphrasing tools in academic writing in terms of strengths, weaknesses, and effectiveness in rephrasing. *Current Trends in Computing, 3*(1), pp. 44 - 57. doi:https://doi.org/ 10.71074/CTC.1653566.

Nitzberg, M., & Zysman, J. (2021). Algorithms, Data, and Platforms: The Diverse Chall. *Journal of European Public Policy*, p. 26. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3802088.

Pham, T. N. (2022). Using Artificial Intelligence Applications For Teaching and Learning Foreign Language: Facts and Solutions. *International Journal of English Language Education and Literature Studies*, pp. 12-15. doi:10.22161/ijeel.1.4.2.

Sample Size Essentials for Educational Studies. (2025, May 24). Retrieved from https://www.numberanalytics.com/blog/sample-size-essentials-educational-studies.

Smith, M., & Scott, H. (2023). Distance Education under Oppression: The Case of Palestinian Higher Education. *Education Sciences, 13*(7). doi:https://doi.org/10.3390/educsci13070729.

Stanford Review. (2024). (2024). How disruptive is AI in students' academic writing in higher education? Retrieved from https://stel.pub.

Stock, F., & Ron, Y. (2023). Issues of integration, language and identity among Palestinian students at an Israeli "Hebrew-speaking" University. *Conflict Resolution Quarterly, 41*(1), pp. 45 - 60. doi:https://doi.org/10.1002/crq.21400.

Sun, P., Anbarasan, Z., & Kumar, D. (2020). Design of online intelligent English teaching platform based on artificial intelligence techniques. *Computational Intelligence*, pp. 1166 - 1180. Retrieved from https://doi.org/10.1111/coin.12351.

Sun, S., Schmidt, K., & Henry, T. (2025). Don't let your Likert scales grow up to be visual analog scales: Understanding the relationship between number of response categories and measurement error. Retrieved from https://arxiv.org/pdf/2502.02846.

Sun, Y., Gregory, J., & Polastri, A. (2024). AI hallucination: Towards a comprehensive classification of the phenomenon. *Humanities and Social Sciences Communications*. Retrieved from https://doi.org/10.1057/s41599-024-03811-x.

Usdan, J., Chang, H., & Pensky, A. (2024). *Generative AI significantly boosts graduate-level writing efficiency and quality.* Heinz College Working Paper, Carnegie Mellon University.

Wecks, J., Voshaar, J., & Plate, B. (2024). Generative AI usage and exam performance. Retrieved from https://doi.org/10.48550/arXiv.2404.19699.

Widiati, U., Rusdin, D., Indrawati, I., & Govender, N. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*. Retrieved from https://doi.org/10.1080/2331186X.2023.2236469.

Woo, D., & Guo, K. (2023). Exploring an AI-supported approach to creative writing: Effects on secondary school students' creativity. doi:10.13140/RG.2.2.24489.06247.

Xu, Z. (2025). Patterns and purposes: A cross-journal analysis of AI tool usage in academic writing. Retrieved from https://arxiv.org/abs/2502.00632.

Yagazi, F. H. (2019). Using Artificial Intelligence Applications to Support University Education in Saudi Arabia. *ASEP*, pp. 259-282.

Zhang, L., & Wu, Z. (2024). Enhancing Postgraduate Academic Writing Skills Through Course Reform: An Action Research Study. *Social Education Research*, pp. 217–233. Retrieved from https://doi.org/10.37256/ser.5220244202.

Zhang, R. (2022). Research on innovative strategies of college students' English teaching under the blessing of big data. *Applied Mathematics and Nonlinear Sciences*, pp. 75 - 84. Retrieved from https://doi.org/10.2478/amns.2021.2.00227.

Zhang, R. (2022). Research on innovative strategies of college students' English teaching under the blessing of big data. *Applied Mathematics and Nonlinear Sciences*, pp. 75 - 84. Retrieved from https://doi.org/10.2478/amns.2021.2.00227.

Appendices

Appendix (A)

Questionnaire

An-Najah National University

Faculty of Graduate Studies

Dear respondent,

The researcher is conducting a study entitled " **The Effect of AI on English Language Research Quality of Palestinian Graduate Students**" as a requirement for obtaining a Master's degree in Methods of Teaching English from An-Najah National University. To achieve the objectives of this study, the researcher developed the attached questionnaire, which has been developed under the supervision of Dr. Nedal Jayousi.

Your information will be confidential and the results of this study will be used only for the purposes of the study.

* If you have any further questions about the questionnaire, please do not hesitate to contact me at amaniabed1241995@gmail.com

Researcher: Amani Abed

Section One: Personal Information

Gender:	☐ Male	☐ Female	
Age:	☐ 22-30	☐ 31-39	☐ 40-48
Level of competence in Information Technology.	☐ Low	☐ Medium	☐ High
Level of competence in Artificial Intelligence (AI)	☐ Low	☐ Medium	☐ High

Section Two:

Kindly respond to the following statements placing (√) next to the answer that corresponds to your response.

#	Statement	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
First domain: Assessing the Effect of AI Tools on Research Quality and Efficiency						
1	AI tools improve the quality of my academic research.					
2	Using AI tools saves my time during the scientific research process.					
3	AI tools help me achieve more accurate results in my research.					
4	AI tools make it easier for me to collect data.					
5	AI tools enhance the organization and structure of my research.					
6	AI tools make it easier for me to analyze data.					
7	The cost of accessing advanced AI tools limits my ability to use them effectively in my research.					
8	I find it difficult to access suitable AI tools for my research needs.					
	Second domain: Identifying Strengths and Weaknesses of AI Tools					
1	**Recall (Strength)** AI tools like Grammarly and Turnitin help me correct grammar and spelling errors in my academic writing.					
2	**Comprehension (Strength)** Using AI tools improves the clarity and structure of my sentences in English research papers.					
3	**Application (Strength)** AI tools save my time by automating repetitive tasks like proofreading and formatting references.					
4	**Synthesis (Strength)**					

	AI tools provide suggestions for improving the vocabulary and terminology used in my research.					
5	**Analysis (Weakness)** AI tools sometimes fail to understand the cultural or academic context of my research topic.					
6	**Evaluation (Weakness)** AI-generated content often lacks originality and depth compared to human-written text.					
7	**Application (Weakness)** I find it difficult to rely on AI tools for complex tasks like analyzing data or interpreting results.					
8	**Analysis (Weakness)** AI tools can produce inaccurate or irrelevant suggestions that require manual correction.					
Third domain: Strategies to Maximize the Benefits of AI Tools						
1	Participating in training sessions on AI tools would enhance my ability to effectively utilize these tools for improving the quality of my English language research.					
2	Collaborating with peers or professors helps me use AI tools more effectively.					
3	Having a user guide for AI tools would be beneficial for my research.					
4	Selecting AI tools that match my research topic improves outcomes.					
5	Using AI tools strategically at different stages of research (e.g., planning, drafting, revising) helps me maximize their benefits without becoming overly dependent on them.					
6	Using AI tools alongside traditional research methods (e.g., brainstorming, peer feedback) helps me maintain my critical thinking and creativity in research.					
7	Implementing strategies like cross-checking AI outputs with academic guidelines helps me overcome ethical concerns about AI-generated content.					
8	The lack of transparency in how AI tools generate results undermines thei reliability for my research.					

Thank You

جـامعـــة النجاح الوطنية
كليـــة الدراســـات العليـــا

تأثير الذكاء الاصطناعي على جودة البحث في اللغة الإنجليزية لدى طلاب الدراسات العليا الفلسطينيين

إعداد
اماني عدنان نعيم عابد

إشراف
د. نضال جيوسي

قدمت هذه الرسالة استكمالاً لمتطلبات درجة الماجستير في أساليب تدريس اللغة الإنجليزية في كلية الدراسات العليا في جامعة النجاح الوطنية في نابلس، فلسطين.
2025

تأثير الذكاء الاصطناعي على جودة البحث في اللغة الإنجليزية لدى طلاب الدراسات العليا الفلسطينيين

إعداد

اماني عدنان نعيم عابد

إشراف

د. نضال جيوسي

الملخص

هدفت هذه الدراسة إلى فحص أثر أدوات الذكاء الاصطناعي على جودة البحوث العلمية في اللغة الإنجليزية لدى طلبة الدراسات العليا الفلسطينيين. وجاءت هذه الدراسة استجابةً للانتشار السريع لتقنيات الذكاء الاصطناعي في المجال الأكاديمي، وقلة الأدلة المحلية المتعلقة بأثرها على ممارسات البحث. وقد سعت الدراسة على وجه التحديد إلى استكشاف تصورات الطلبة تجاه دور الذكاء الاصطناعي في تحسين جودة البحث، والتعرف إلى نقاط القوة والضعف في هذه الأدوات، إضافة إلى الاستراتيجيات التي يمكن أن تدعم استخدامها الفعّال.

اتبعت الدراسة المنهج الوصفي الكمي، واعتمدت على استبانة وُزِّعت إلكترونيًا بسبب القيود الناتجة عن الحرب في غزة وصعوبات الحركة في الضفة الغربية. بلغ حجم المجتمع المستهدف 1931 طالبًا وطالبة من برامج الدراسات العليا، وكان من المخطط الحصول على عينة مقدارها 320. إلا أن عدد الاستبانات الصحيحة التي جُمعت بلغ 171، أي ما نسبته 8.9% من المجتمع. تضمنت الاستبانة 24 فقرة موزعة على ثلاثة مجالات، وقد أُكِّد صدقها من خلال لجنة من الخبراء، بينما بلغت معاملاتها الخاصة بالثبات (كرونباخ ألفا) 0.907، وهو ما يشير إلى مستوى عالٍ من الموثوقية.

أظهرت النتائج أن الطلبة عبّروا عمومًا عن تصورات إيجابية تجاه أدوات الذكاء الاصطناعي، حيث بلغ المتوسط الكلي عبر المجالات الثلاثة 3.58 (71.6%)، وهو ما يُصنّف بدرجة عالية. وقد قيّم الطلبة بشكل

خاص قدرة هذه الأدوات على توفير الوقت، وتنظيم المعلومات، وتحسين الدقة، إذ سُجّل أعلى متوسط في مجال الاستراتيجيات (M = 3.62 ، 72.4% .(وفي المقابل، برزت بعض المخاوف المتعلقة بضعف الأصالة، وصعوبة فهم السياقات الأكاديمية أو الثقافية، إضافة إلى تحديات تتعلق بالوصول إلى الأدوات المتقدمة. كما لم تُظهر النتائج فروقًا دالة وفقًا لمتغيرات الجنس أو العمر أو المعرفة التقنية أو الخبرة السابقة بالذكاء الاصطناعي.

يوصى بأن تعمل الجامعات الفلسطينية على تنظيم ورش تدريبية تمكّن طلبة الدراسات العليا من استخدام أدوات الذكاء الاصطناعي بفاعلية وأخلاقيات عالية. كما ينبغي دمج الوعي بتطبيقات الذكاء الاصطناعي في مساقات مناهج البحث العلمي، إلى جانب توفير أدوات موثوقة يسهل الوصول إليها. ويوصى أعضاء هيئة التدريس بتشجيع الطلبة على تحقيق التوازن بين الاستعانة بتلك الأدوات وممارسة التفكير النقدي والأصالة البحثية.

الكلمات المفتاحية: الذكاء الاصطناعي، جودة البحث، طلبة الدراسات العليا، الكتابة الأكاديمية، فلسطين