



REPUBLIC OF TÜRKİYE

ÇANAKKALE ONSEKİZ MART UNIVERSITY

SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF FOREIGN LANGUAGE EDUCATION

**ETHICAL IMPLEMENTATION OF ARTIFICIAL INTELLIGENCE
IN THE PROCESS OF ACADEMIC WRITING**

MASTER'S THESIS

BURCU ÖZGE RAZI

Supervisor

PROF. DR. AYSUN YAVUZ

ÇANAKKALE – 2024



REPUBLIC OF TÜRKİYE

ÇANAKKALE ONSEKİZ MART UNIVERSITY

SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF FOREIGN LANGUAGE EDUCATION

**ETHICAL IMPLEMENTATION OF ARTIFICIAL INTELLIGENCE IN THE
PROCESS OF ACADEMIC WRITING**

MASTER'S THESIS

BURCU ÖZGE RAZI

Supervisor

PROF. DR. AYSUN YAVUZ

ÇANAKKALE – 2024



T.C.
ÇANAKKALE ONSEKİZ MART ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ



Burcu Özge RAZI tarafından Prof. Dr. Aysun YAVUZ yönetiminde hazırlanan ve 28/08/2024 tarihinde aşağıdaki jüri karşısında sunulan “**Ethical Implementation of Artificial Intelligence in the Process of Academic Writing**” başlıklı çalışma, Çanakkale Onsekiz Mart Üniversitesi Lisansüstü Eğitim Enstitüsü **Yabancı Diller Eğitimi Anabilim Dalı**’nda **YÜKSEK LİSANS TEZİ** olarak oy birliği ile kabul edilmiştir.

Jüri Üyeleri

İmza

Prof. Dr. Aysun YAVUZ
(Danışman)

Doç. Dr. Kutay UZUN

Dr. Öğr. Mustafa TEKİN

.....

.....

.....

Tez No : 10608440

Tez Savunma Tarihi : 28/08/2024

.....
Doç. Dr. Melis ULU DOĞRU
Enstitü Müdürü

28/08/2024

DECLARATION OF ETHICS

I hereby declare that this thesis, which I prepared in accordance with the Writing Guidelines of the School of Graduate Studies of Çanakkale Onsekiz Mart University, adheres to academic and ethical standards. I have obtained all data, information, and documents within these guidelines, and presented them with scientific integrity. All sources used are appropriately cited, and no alterations have been made to the data. This work is entirely original. Should there be any discrepancies, I accept full responsibility for any resulting consequences.

ETİK BEYAN

Çanakkale Onsekiz Mart Üniversitesi Lisansüstü Eğitim Enstitüsü Tez Yazım Kuralları'na uygun olarak hazırladığım bu tez çalışmasında; tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi, tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu, tez çalışmasında yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi, kullanılan verilerde herhangi bir değişiklik yapmadığımı, bu tezde sunduğum çalışmanın özgün olduğunu, bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi taahhüt ve beyan ederim.

Burcu Özge RAZI

28/08/2024

ACKNOWLEDGEMENTS

There are many people who deserve my deepest thanks and appreciation for their unwavering support and belief in me throughout this journey. Words alone cannot express my gratitude for their contributions, motivation, and encouragement.

First, I would like to express my profound gratitude to my supervisor, Prof. Dr. Aysun Yavuz, for her exceptional guidance and motivation. Her extensive knowledge and invaluable suggestions were instrumental in the completion of this study. Her unwavering support and timely advice provided me with the confidence to persevere.

I am also deeply grateful to the participants and the instructor of the course whose permission to collect data in his class significantly contributed to this research. His cooperation and understanding were invaluable.

I extend my heartfelt thanks to my thesis committee members, Dr. Mustafa Tekin, and Assoc. Prof. Dr. Kutay Uzun, for their constructive feedback and insightful contributions to my research. Their expertise and support were crucial in shaping the final outcome of this study.

A special thanks goes to my close friend Beste Sezen for her emotional and academic support throughout this process. Her assistance in providing interrater reliability was crucial, and her unwavering belief in me was a constant source of encouragement. Additionally, I am grateful to her mother, Emine Sezen, for taking care of my daughter Ece with the tenderness of a grandmother, allowing me to focus on my work.

I would like to thank Emine Özer for being a safe haven for me and for treating me like a sister. Thank you for being here Emine Hocam!

I am deeply indebted to Neslihan Saçan Gözütok, my dearest friend and confidant. Words cannot express how much her friendship, support, and understanding have meant to me. She is my rock, my partner in crime, my unbiological sister and I love her dearly.

Ve kıymetli ailem, kız kardeşim Duygu Karataş Boudrioua, annem Hamide Karataş ve babam Fikret Karataş'a teşekkür etmek istiyorum. Annemin bana verdiği destek ve kızım Ece'ye mükemmel bir anneanne oluşu paha biçilemez. Kız kardeşim Duygu, bunaldığım zamanlarda destekçim, sırdaşım ve sığınacak limanım oldu. Ablam Duygu Karataş Boudrioua, Dünyanın en mükemmel annesi olarak her gün bana ilham veriyor. Ailemin sabrı, ilgisi ve kesintisiz desteği için minnettarım.

İkinci ailem Fahriye ve Ahmet Razi, Duygu Üzmezoğlu Razi, Cüneyd, Mert ve Emre Üzmezoğlu'na da kocaman bir teşekkür borçluyum. Güzel kızımın babaannesi Fahriye Annemin gösterdiği sevgi ve destek benim için çok değerliydi. Kendisi mükemmel bir babaanne ve ben çalışırken sayesinde gözüm hiç arkada kalmadı. Dedesi Ahmet Razi, Ece'nin en güzel oyun arkadaşı oldu. Abileri Mert ve Emre ise Ece'den sevgilerini hiç esirgemedi ve her şeylerini onunla paylaştılar. Sevgili halası Duygu ve eniştesi Cüneyd, Ece için her zaman sevgi dolu limanlar oldular. İyi ki varsınız!

Finally, I want to express my deepest love and gratitude to my dear husband, Salim Razi. Words cannot capture the depth of my love for him. He is a wonderful father, a perfect husband, and the most patient and caring person I know. His support, patience, and willingness to share the responsibilities of our home and child have been my greatest strength. Salim, you are my greatest blessing and the best part of my life. I am incredibly fortunate to have you by my side. If during this time I have wronged you, hurt you, or upset you, please forgive me. I love you so much.

I dedicate this Master's Thesis to my beloved daughter Ece and my precious niece İda and all the babies and children who lost their lives in Palestine...

Burcu Özge RAZI
Çanakkale, Ağustos 2024

ÖZET

YAPAY ZEKANIN AKADEMİK YAZIM SÜRECİNDE ETİK KULLANIMI

Burcu Özge RAZI

Çanakkale Onsekiz Mart Üniversitesi

Lisansüstü Eğitim Enstitüsü

Yabancı Diller Eğitimi Anabilim Dalı Yüksek Lisans Tezi

Danışman: Prof. Dr. Aysun YAVUZ

28/08/2024, 159

Çalışmada, İngilizce öğretmen adayları için akademik yazım sürecinde yapay zekanın (YZ) etik kullanımını araştırmaktadır. Araştırma, araçsal kolektif vaka çalışmasıyla karma yöntem yaklaşımını kullanmıştır. Çalışma, öğretmen ve akran dönütüyle karşılaştırıldığında YZ tarafından üretilen dönütün Anonim Çok Aracılı Yazma Modeli aracılığıyla etkinliğini incelemektedir. Çalışma, Türkiye'deki bir üniversiteden 16 öğrenciyi içermektedir ve bu öğrenciler öğretmenlerden, akranlardan ve ChatGPT'den akademik makalelerine dönüt almışlardır. Nicel veriler, dönüt öncesi ve sonrası makale puanlarıyla toplanırken, nitel içgörüler katılımcılarla yapılan bireysel görüşmeler aracılığıyla elde edilmiştir. Bulgular, öğretmen dönütünün öğrenci performansında en yüksek iyileşmeyi sağladığını göstermektedir. Öğretmenlerin verdiği puanlar ile ChatGPT'nin verdiği puanlar arasında anlamlı bir fark bulunmamıştır. Ancak, öğretmenlerin verdiği puanlar ile akranların verdiği puanlar arasında anlamlı bir fark vardır ve akran dönütü puanlarının öğretmeninkilerden daha yüksektir. YZ tarafından üretilen ve akran dönütü orta düzeyde iyileşmeler göstermiş, YZ dönütü akran dönütüne yakın bir etkililik sergilemiştir. Nitel veri bulgularına göre, öğrenciler YZ'nin sağladığı anında ve ayrıntılı dönütü, belirli geliştirme alanlarını belirlemelerine yardımcı olduğu için beğenmişlerdir. Ancak, YZ dönütünün insana benzemeyen doğası ve teknolojiye aşırı bağımlılık potansiyeli hakkında da endişeler dile getirilmiştir. Akran dönütü işbirlikçi yönüyle değerli bulunmuş, ancak akran yeterlilik düzeylerinin değişkenliği nedeniyle bazen daha az güvenilir olarak görülmüştür. Öğretmen dönütü en güvenilir ve kapsamlı bulunmuş, eleştiri ve teşvik dengesini sağlamıştır. Teknik sorunlar ve akademik bütünlük konusundaki endişeler gibi zorluklar da nitel veriler

sonucunda ortaya çıkmıştır. Çalışma, YZ'nin eğitim uygulamalarına entegrasyonu için etik yönergeler ve öğretmen eğitimi gerekliliğini vurgulayarak önerilerle son bulmaktadır.

Anahtar Kelimeler: Akademik yazma, İngiliz dili eğitimi, Üretken yapay zeka, akademik dürüstlük, Yapay zekanın etik kullanımı, Dönüt



ABSTRACT

ETHICAL IMPLEMENTATION OF ARTIFICIAL INTELLIGENCE IN THE PROCESS OF ACADEMIC INTEGRITY

Burcu Özge RAZI

Çanakkale Onsekiz Mart University

School of Graduate Studies

Department of Foreign Languages Education English Master's Thesis

Supervisor: Prof. Dr. Aysun YAVUZ

28/08/2024, 159

The present study investigates ethical uses of artificial intelligence (AI) in academic writing in foreign language education for prospective teachers. Utilizing mixed-method approach with instrumental collective case study design, this research examines effectiveness of AI-generated feedback compared to traditional teacher and peer feedback through the Anonymous Multi-Mediated Writing Model. The study involves 16 students from a university in Türkiye, who received feedback on their academic essays from teachers, peers, and ChatGPT. Quantitative data were collected through pre- and post-feedback essay scores, while qualitative data were gathered via individual interviews. The findings reveal that teacher feedback made the highest improvement in student performance. There was no significant difference between the scores by teachers and by ChatGPT. However, there was a significant difference between the scores given by teachers and peers, with peer scores being higher than those of the teacher. Both AI-generated and peer feedback showed moderate improvements, with AI feedback close to the effectiveness of peer feedback. Regarding qualitative data findings, students appreciated immediate and detailed feedback provided by AI. However, they also expressed concerns about impersonal nature of AI feedback and the potential for over-reliance on technology. Peer feedback was valued for collaborative aspect, though it was sometimes seen as less reliable. Teacher feedback was considered the most trustworthy and comprehensive. The study concludes with recommendations for educational practices, ethical guidelines and teacher training to harness AI's potential while maintaining academic standards.

Keywords: Academic writing, English language teaching, generative artificial intelligence, academic integrity, ethical use of artificial intelligence, feedback



TABLE OF CONTENTS

	Page No
APPROVAL	i
ETHICAL DECLARATION	ii
ACKNOWLEDGEMENTS	iii
ÖZET	v
ABSTRACT	vii
TABLE OF CONTENTS	ix
LIST OF ABBREVIATIONS	xv
LIST OF TABLES	xvi
LIST OF FIGURES	xvii

CHAPTER ONE

INTRODUCTION	1
--------------	---

1.1. Background of the Study	1
1.2. Problem Statement	5
1.3. The Purpose of the Study	6
1.4. Research Questions	7
1.5. The Significance of the Study	7
1.6. The Definition of Key Terms	8
1.7. Limitations of the Study	9

1.8. The Summary of the Chapter	9
---------------------------------------	---

CHAPTER TWO

LITERATURE REVIEW

	11
2.1. Teaching Writing	11
2.1.1. Approaches to Teaching Writing	12
Product Writing	12
Process Writing	13
Genre Approach	13
Genre Process Approach	14
2.1.2. Sources of Feedback	16
Teacher Feedback	16
Peer Feedback	20
2.1.3. Ways of Giving Feedback	22
Face-to-face Feedback	22
Online Feedback	24
2.1.4. Types of Feedback	25
2.2. Anonymous-Multi Mediated Writing Model	27
2.2.1. Study by Razi in 2016	28
2.2.2. Study by Razi in 2017	29
2.2.3. Study by Razi in 2023	30
2.2.4. Summary of Razi's Selected Studies	31
2.2.5. Processes of Anonymous-Multi Mediated Writing Model	32

2.2.6.	Theoretical Background of Anonymous-Multi Mediated Writing Model	33
2.3.	Generative Artificial Intelligence	34
2.3.1.	Prompt Engineering	36
2.3.2.	Generative Artificial Intelligence Research in Education	37
2.3.3.	Ethical Considerations for the Use of Generative Artificial Intelligence in Writing	42
2.3.4.	Generative Artificial Intelligence based Writing	44
2.4.	Similar Studies to the Current Study	47
2.5.	The Summary of the Chapter	50
 CHAPTER THREE METHODOLOGY		52
3.1.	Research Design	52
3.2	Research Questions	54
3.3.	Setting and Participants	54
3.3.1.	Structure of the Course	55
3.3.2.	Selection of Participants	57
3.4.	Data Collection Tools and Data Collection Sources	59
3.4.1.	Fundamentals of Academic Writing Rubric as a Data Collection Tool ..	60
3.4.2.	Assignments of the Course as Data Collection Sources	61
3.4.3.	Development of the Prompts as Data Collection Sources	62
3.4.4.	Development of Interview Questions as a Data Collection Tool	66
3.5.	Data Collection Procedures	67

3.5.1. Data Collection of the Assignment 1 and the Teacher Feedback	67
3.5.2. Data Collection of the Assignment 2 and Peer Feedback	67
3.5.3. Data Collection of the Assignment 3 and GenAI Feedback	68
3.5.4. Interviews with the Participants	69
3.6. Data Analysis	69
3.6.1. Analysis Procedures for Research Question 1	69
3.6.2. Analysis Procedures for Research Question 2	72
3.6.3. Analysis Procedures for Research Question 3	73
3.6.4. Inter-rater Reliability	74
3.7. The Summary of the Chapter	75
 CHAPTER FOUR FINDINGS 	
4.1. Comparative Effect of Feedback Provided by the Teacher, Peers, and Generative Artificial Intelligence	77
4.1.1. The Difference between the Gain Scores of the Feedback Provided by Teachers, Peers, and Generative Artificial Intelligence	77
4.1.2. The Difference between the Pre-scores given by the Teacher and the Scores Given by Peers and Generative Artificial Intelligence	79
4.2. The Effects of Feedback from the Teacher, Peers, and Generative Artificial Intelligence on the Scores of the Assignments	81
4.2.1. The Effect of Teacher Feedback	81
4.2.2. The Effect of Peer Feedback	87
4.2.3. The Effect of Generative Artificial Intelligence Feedback	95
4.2.4. The Impact of Feedback on Rubric Elements across Assignments	101

4.3.	Strengths and Weaknesses of Feedback from the Teacher, Peers, and Generative Artificial Intelligence	104
4.3.1.	Strengths and Weaknesses of Teacher Feedback	104
4.3.2.	Strengths and Weaknesses of Peer Feedback	109
4.3.3.	Strengths and Weaknesses of Generative Artificial Intelligence Feedback	114
4.3.4.	Themes Emerged as the Weaknesses and Strengths of Feedback from the Teacher, Peers, and Generative Artificial Intelligence	119
4.4.	The Summary of the Chapter	123

CHAPTER FIVE

	DISCUSSION, CONCLUSION, AND IMPLICATIONS	125
--	--	-----

5.1.	Discussion of the Findings	125
5.1.1.	Discussion of the Research Question 1	125
	Research Question 1.1.	125
	Research Question 1.2.	127
5.1.2.	Discussion of the Research Question 2	129
5.1.3.	Discussion of the Research Question 3	131
5.2.	Conclusions	133
5.2.1.	Conclusions from the Research Question 1	134
5.2.2.	Conclusions from the Research Question 2	135
5.2.3.	Conclusions from the Research Question 3	137
5.2.4.	Ethical Use of Generative Artificial Intelligence in the Process of Academic Writing	139
5.3.	Implications	139

5.3.1. Educational Implications	139
5.3.2. Policy Implications	140
5.3.3. Student Implications	141
5.3.4 Academic Integrity Implications	142
5.4. Suggestions for Further Research	142
5.5. The Summary of the Chapter	143
REFERENCES	145
APPENDICES	I
APPENDIX A	I
APPENDIX B	IX
APPENDIX C	XII
APPENDIX D	XIX
APPENDIX E.....	XX

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
AMMW	Anonymous Multi-Mediated Writing
APA	American Psychological Association
AWE	Automated Writing Evaluation
CALL	Computer-Assisted Language Learning
CHATGPT	Chat Generative Pre-Trained Transformer
CF	Corrective Feedback
COPE	Committee on Publication Ethics
COVID	Corona Virus Disease
EFL	English as a Foreign Language
ELT	English Language Teaching
ENL	English as a New Language
ESL	English as a Second Language
ESP	English for Specific Purposes
FAWR	Fundamentals of Academic Writing Rubric
IBM	International Business Machines Corporation
ICAI	International Center for Academic Integrity
ICT	Information and Communication Technology
L1	First Language
L2	Second Language
LLM	Large Language Models
SFL	Systemic Functional Linguistics
SPSS	Statistical Package for the Social Sciences
TAWR	Transparent Academic Writing Rubric
WRF	Written Corrective Feedback
ZPD	Zone of Proximal Development

LIST OF TABLES

Table No	Table	Page No
Table 1	Comparisons of Participants' Scores with Mean, Minimum, and Maximum Values	58
Table 2	Key Components of the Research Methodology	60
Table 3	Wilcoxon Signed Ranks Test Results for Comparison of Gain Scores Across Different Feedback Types for Essay Assignments	79
Table 4	Wilcoxon Signed Ranks Test Results for Comparison of Pre-scores across Different Feedback Types for Essay Assignments	81
Table 5	Pre- and Post-Score Mean Values for 20 Rubric Elements Across Three Assignments	102
Table 6	The Frequencies of the Strengths and Weaknesses Identified for Teacher, Peer, and GenAI Feedback	120

LIST OF FIGURES

Figure No	Figure	Page No
Figure 1	Grading Structure Diagram for the Writing Skills I Course	57
Figure 2	Participant Selection Process	59
Figure 3	Development of the Prompt – Stage 1 – Getting Ready for the Context Manager Pattern	63
Figure 4	Formation of the Prompt – Stage 2 – Formation Of The Prompt With The Context Manager Pattern	64
Figure 5	Question Refinement Pattern – Stage 3	64
Figure 6	Custom Instructions for ChatGPT	65
Figure 7	Data Analysis Procedures for Research Question 1.1.	70
Figure 8	Data Analysis Procedures for Research Question 1.2.	71
Figure 9	Data Analysis Procedures for Research Question 2	74

CHAPTER ONE

INTRODUCTION

In this chapter, nine sections were covered including background of the study, the problem statement, aims of the study, research questions, significance of the study, the scope of the study, limitations of the study, and related terminology. Lastly, the chapter was concluded with the chapter summary.

1.1. Background of the Study

The realm of foreign language teaching has undergone significant paradigm shifts over time. The analysis of grammar and rhetoric in classic Latin literature laid the foundation for foreign language education between the 17th and 18th centuries. Consequently, modern languages entered the European curriculum in the 18th century and were taught using the Grammar-Translation Method (Richards and Rodgers, 2001). These language-centred methods eventually gave way to learner-centred approaches that emphasized interaction among learners (Kumaravadivelu, 2006) since opposition arose regarding the use of the Grammar-Translation method as a tedious and unengaging experience (Richards and Rodgers, 2001). This shift marked the transition from rigid, grammar-focused teaching to more interactive and student-centred methodologies.

As the 19th century began, the reform movement advocated for naturalistic approaches to language learning, leading to the emergence of the Natural Method (Richards and Rodgers, 2001). Moving into the 20th century, various methods such as the Audiolingual Method, Situational Method, Communicative Approach, Silent Way, Natural Approach, Total Physical Response, Content-Based Instruction, Task-Based Language Teaching, Competency-Based, Cooperative Learning, and Whole Language Approach emerged in succession (Richards and Rodgers, 2001). However, the end of the 20th century signalled another paradigm shift, pioneering what linguists refer to as the Post-methods era. In short, this period reflected changing educational philosophies and approaches, and was finalized in the current various number of methodologies.

The transition to the Post-Methods era in language teaching is supported by several arguments from various linguists including Richards and Rodgers, Brown, Kumaravadivelu and many more. For example, Richards and Rodgers (2001) claimed that traditional methods were no longer central to effective language teaching, and Brown (2000) argued that methods were too prescriptive and overgeneralized; thus, they failed to account for the complexity of teaching contexts and often served the interests of powerful centres, promoted linguistic imperialism. Brown (2000) maintained this idea by stating that ideas and methods lack empirical validation and fail to address broader teaching and learning contexts. Building on this statement, Richards and Rodgers (2001) emphasized that good teaching was not just about following a method's principles, as this could make learners passive and make teachers mere implementers of prescribed practices. They highlighted the importance of contextual factors and comprehensive curriculum development, and noted that many methods lack a solid research foundation (Kumaravadivelu, 2008). Furthermore, Kumaravadivelu critiqued methods for being based on idealized scenarios that could not account for the diverse needs of real-world teaching, and underscored that their inadequacy in addressing the complexities of language teaching globally. As a conclusion, these critiques highlighted the need for more flexible and context-sensitive approaches in the Post-Methods era.

In addition to traditional pedagogical approaches such as the Grammar-Translation method, the integration of technology into foreign language education started with audio recordings, mainframe computers, language laboratories, and interactive computer-based learning methods (Richards and Rodgers, 2001). The advent of mainframe computers, which allowed for multiple terminals, paved the way for Computer-Assisted Language Learning (CALL) (Hubbard, 2017). Towards the end of the 20th century, several innovative projects were conducted, including Programmed Logic for Automated Teaching Operations, the Tutorial Russian project at Stanford University, and the Time-shared Interactive Computer Controlled Information Television, which offered possibilities for providing feedback and saving records (Hubbard, 2021). Throughout the early 21st century, CALL continued to evolve from being limited by time and place to becoming ubiquitous (Hubbard, 2021), with the World Wide Web marking an era characterized by the continual enhancement and evolution of web-based educational tools (Hubbard, 2021). The integration of technology significantly transformed the landscape of language education.

In the concluding remarks of his book ‘An Invitation to CALL,’ Hubbard (2021) noted, “One thing we know – artificial intelligence (AI) is in our future” (p. 78). Indeed, in November 2022, OpenAI launched ChatGPT, exemplifying the significant strides made in AI within the realm of language learning. The evolution of technology in language education has moved from basic tools to sophisticated AI applications, transforming the way languages are taught and learned. This progression underscores the impact of AI in revolutionizing language learning methodologies.

The introduction of AI technology was met with much bias and criticism. A significant portion of the literature on artificial intelligence consisted of concerns about academic integrity and the restrictions on the use of AI (Sullivan et al., 2023), provoking suspicions of teachers regarding their students' involvement in several types of academic misconduct, including plagiarism (Cotton et al., 2024; Moya et al., 2024). However, while it is obligatory to address this issue and discuss detective or reactive methods, such a singular focus may leave us inactive in the face of this ever-evolving technology. Since AI continuously improves, avoidance, denial, or dismissal of this rapidly advancing technology is not a viable option. Numerous guidelines for the ethical use of artificial intelligence in both theory and practice have been disseminated by universities, organizations, and academic integrity networks such as The European Commission (2022) and the International Center for Academic Integrity (ICAI, n.d.). Addressing these concerns proactively can help harness AI’s potential while maintaining academic integrity.

Expert scholars on artificial intelligence and education have stated that AI benefits both students and teachers. Eaton (2023) highlighted the need for transdisciplinary research on academic integrity and artificial intelligence, focusing on the responsible, equal, ethical, and inclusive use of AI. Eaton noted that AI could boost human creativity, overcome language barriers, and create a hybrid writing style of AI and human input under human responsibility, with additional attribution to the AI. Consequently, a new paradigm shift has emerged. As Eaton (2023, p. 3) stated, “Postplagiarism refers to an era in human society in which advanced technologies, including artificial intelligence and neurotechnology, including brain-computer interfaces (BCIs), are a normal part of life, including how we

teach, learn, and interact daily.” This highlights the potential for AI to revolutionize educational practices while addressing ethical considerations.

Over two decades ago, Prensky (2001) introduced the term ‘digital natives’ to describe individuals born into a world of technological marvels, naturally used to the advancements of the era, contrary to digital immigrants who adapt to the advancements. Today, it is apt to say that there are digital natives born into the emergence of artificial intelligence. These digital natives of artificial intelligence will grow up with AI as a fundamental aspect of their environment. Meanwhile, digital immigrants of artificial intelligence have found and will continue to find themselves navigating the winds of change created by these AI-native generations. To thrive in this dynamic environment, immigrants must adapt to the language of AI and embrace these technological advancements. While winds can be disruptive, they also carry the potential for growth like the seeds of flowers dispersed by these winds. This metaphor illustrates the transformative potential of AI in education by resembling it to seeds carried by the wind that can bloom into something beautiful under the right conditions.

As a student, encountering restrictions like ‘the use of artificial intelligence in assignments is strictly prohibited’ is disheartening for me. I regard AI as a seed carried by the winds of technological change, and it is the one that has the potential to bloom into something beautiful. However, for this growth to occur, it requires the right conditions including the transparent use of AI within ethical frameworks. I need to understand how AI can enrich my imagination, enhance my performance, and contribute to my critical stance. In essence, there are vibrant flower gardens where the seeds of artificial intelligence which were brought by the wind of technological change can flourish. Thus, if AI is used ethically, can significantly enhance the learning experience.

Being a student wandering around the digital landscape, I have sown the seeds of change by delving into the realm of artificial intelligence for my master’s thesis, and I focused on the ethical use of AI in academic writing. My goal is to cultivate the seed I have planted into a flower of responsible AI use. I believe my peers are also nurturing their seeds,

collectively sowing gardens that will shape the ethical use of artificial intelligence for future generations of digital natives. Together, we will embrace the ethical horizon and hopefully witness the growth of AI as a valuable ally in our academic pursuits. This collective effort emphasizes the importance of responsibly integrating AI into education for the benefit of future generations.

1.2. Problem Statement

The growth of artificial intelligence-based text generation technologies has introduced a range of complex challenges to academic integrity, which require careful consideration within the context of academic institutions and their policies (Kumar et al., 2023). With the advent of AI-powered text generation tools like ChatGPT, students now may produce papers, essays, or assignments rapidly and with minimal effort, causing concerned academic integrity advocates and academic integrity aware teachers to become more concerned about academic integrity (Kumar et al., 2023). This issue carries the risk of students bypassing ethical writing practices and failing to provide proper attribution (Lancaster, 2023). Thus, students are unable to achieve the outcomes that their teachers aimed for them and ultimately engaged in academic misconduct since it was stated that the overusing of AI tools harms critical abilities of the students, including reasoning, writing, and comprehending (Rodrigues et al., 2023). It is crucial for educational institutions and educators to remain cautious in addressing these challenges posed by AI technology in the realm of education.

Although the use of artificial intelligence is like a seed waiting to sprout, unfortunately, some seeds have fallen into unsuitable places, which is quite normal in nature. If these fallen seeds do not receive sunlight or water, they will not germinate. Similarly, in the context of using artificial intelligence, there are many sources of motivation for users to achieve something through its use. This is actually one of the biggest problems to address regarding the use of artificial intelligence. There are many factors that make the use of artificial intelligence attractive. As expressed by Kumar et al. (2023), some factors include the ability of AI tools to generate unique text each time, making their use complex to

understand, while the benefits of using AI-generated outputs include passing a course, obtaining a degree, saving energy, and achieving financial and emotional gains. At this point, the problems are quite clear. It is, of course, impossible to completely eliminate academic misconduct, and the study does not aim to solve such a problem. Therefore, rather than vilifying artificial intelligence, there is a need for studies that embrace AI and demonstrate how to move forward hand in hand with it.

1.3. The Purpose of the Study

The umbrella aim of this study is to base the use of artificial intelligence, particularly in foreign language education and for prospective foreign language teachers, on educative and pedagogical approaches rather than punitive approaches, as recommended by Bertram Gallant (2008) and by Bretag et al. (2011) in exemplar academic integrity policies. This study primarily aims to demonstrate a sample ethical use of artificial intelligence (AI) that could be possible for both students and teachers by employing AI as a facilitator. Specifically, it seeks to show that AI can be used as a facilitator to provide feedback on academic essays written by foreign language teaching students within the scope of academic writing courses. To achieve this, the study will initially focus on identifying which type of feedback including peer, teacher, or AI most significantly improves student performance. The study will also explore the consistency and reliability of scores assigned by different evaluators, aiming to highlight any discrepancies or agreements between the scores given by teachers, peers, and AI. Thus, it also aims to shed light on whether it is possible to use AI for assessment or not. Additionally, the study aims to understand student perceptions of the feedback they receive by gathering insights from students about the strengths and weaknesses of feedback from teachers, peers, and AI.

1.4. Research Questions

1. Is there a significant difference between:

1.1. the gain scores of feedback provided by teachers, peers, and ChatGPT on students' essay scores?

1.2. the pre-scores given by the teacher and the scores given by peers and Generative artificial intelligence (GenAI)?

2. How does the feedback provided by the teacher, peers, and artificial intelligence affect the improvement of students' assignment grades, as measured by pre- and post-feedback scores?

3. How do students perceive the strengths and weaknesses of feedback provided by teacher, peers, and GenAI?

1.5. The Significance of the Study

The significance of this study lay in its multifaceted implications for several aspects. Firstly, this research aimed to pioneer and promote the use of AI in providing feedback on academic essays, leading to improved student performance and engagement. Thus, this study was significant for being pedagogically innovative. Secondly, it addressed concerns about academic misconduct by shifting the focus from academic misconduct to academic ethical collaboration. In short, this study was significant for providing ethical integration of AI. Thirdly, as the study aimed to compare the consistency of different feedback providers' scores for students' essays, it also had the potential to present different assessment methods for teachers. In summary, it could be timesaving for teachers to integrate AI in their assessment stage. Fourthly, in the age of advanced technologies, this study contributed to the ongoing debate on maintaining academic integrity by supporting the stance of standing with integrity. Lastly, by gathering student perceptions of the feedback they received from teachers, peers, and AI, this study could help educators tailor feedback mechanisms to meet students' needs, thereby designing more effective and engaging learning experiences. In conclusion, the significance of this study was inclusive, covering pedagogical innovation, ethical considerations, and practical implications for educators and students. It aimed to create a foundation for the responsible and effective use of AI in education by contributing to the broader goals of academic integrity and improved learning outcomes.

1.6. The Definition of Key Terms

Academic writing: Academic writing is a type of formal writing in which authors are expected to support their claims by referencing (Razi, 2011).

Generative AI: “Generative AI refers to a set of artificial intelligence techniques and models designed to learn the underlying patterns and structure of a dataset and generate new data points that plausibly could be part of the original dataset” (Pinaya et al., 2023, p. 2).

Feedback: “Feedback as a generic term includes comments/commentary/response as well as corrective feedback (CF), which focuses on formal aspects of learners’ language and is provided with the intent to improve linguistic accuracy. Feedback is often categorized based on the source: teacher-, peer-, or tutor-feedback. It can be provided by the person giving feedback (handwritten or computer mediated) or computer-generated (automatic)” (Ene and Upton, 2018, pp. 1-13).

Academic integrity: “Compliance with ethical and professional principles, standards, practices and consistent system of values, that serves as guidance for making decisions and taking actions in education, research and scholarship” (Tauginienė et al., 2018, p. 7).

Process writing: Process writing makes a reference to a teaching writing approach in which student-authors follow “steps such as prewriting, planning, drafting, and revision” by making use of feedback that they receive before the submission of the final version of the writing assignment (Razi, 2023, p. 2).

Academic misconduct: “Any action or attempted action that undermines academic integrity and may result in an unfair academic advantage or disadvantage for any member of the academic community or wider society” (Tauginienė et al., 2018, p. 8).

English for Academic Purposes: “Language research and instruction that focuses on the communicative needs and practices of individuals working in academic contexts” (Hyland and Shaw, 2016, p. 1).

1.7. Limitations of the Study

The findings of this research should be interpreted within the context of its specific sample and setting, which may restrict their generalizability. The study focused on a particular institution and included a sample size of 16 students and such specificity in both sample and setting can limit the broader applicability of the findings beyond this specific group.

1.8. The Summary of the Chapter

The first chapter of this study traced the evolution of foreign language teaching, and it started with the early Grammar-Translation Method and moved through various approaches, and it was finalised in the modern Post-Methods era. This progression highlighted a shift from rigid, grammar-focused teaching to more interactive and student-centred methodologies, where technology, especially AI tools like ChatGPT, played a crucial role in this transformation.

Following this historical context, the chapter addressed the challenges posed by AI in maintaining academic integrity. However, with AI tools making it easier for students to generate text quickly, concerns about plagiarism and ethical writing practices became more pronounced. Consequently, the study aimed to explore how AI could be used ethically to provide feedback on academic essays by focusing on the effectiveness of feedback from peers, teachers, and AI.

To guide this exploration, the research was structured around questions about the impact of different types of feedback on student performance and about how students

perceived this feedback. The chapter emphasized the significance of the study in promoting the ethical use of AI in education by helping teachers and students find the balance between innovation and integrity.

In addition, by defining key terms like academic writing, generative AI, and academic integrity, the chapter introduced the solid foundation for the study. It also explained the study's limitations with specific emphasis on the findings that were based on a specific sample and setting, which might, consequently, limit their broader applicability. Overall, the chapter underscored the potential for AI to enhance learning when used responsibly and ethically, and it set the stage for the detailed exploration in the subsequent chapters.

CHAPTER TWO

LITERATURE REVIEW

In the second chapter of the study, the literature review was organized into three main themes: teaching writing, GenAI, and prompt engineering. Within the teaching writing theme, the discussion encompassed various approaches such as product writing, genre approach, and process writing. Feedback mechanisms were explored, covering sources, methods, and types of feedback, along with the anonymous multi-mediated writing model. The section on GenAI investigated its evolution, applications in writing, exemplar implementations, potential benefits of its integration, and ethical implications surrounding its deployment. The final focus of the literature review revolved around prompt engineering. The chapter was concluded with a comprehensive summary of the discussed topics.

2.1. Teaching Writing

The Grammar-Translation Method, prominent in the 18th century, emphasized the acquisition of literacy skills through the rigorous analysis of classical texts, fostering a deep understanding of grammar and rhetoric. This historical emphasis on reading and writing classical works laid the groundwork for the modern approach to teaching writing, where the development of literacy skills remains paramount. Although spoken language is acquired naturally, the ability to read and write must be learned intentionally. In the 21st century, it is almost impossible to exist in society without literacy. At this point, education, which requires learners to be proficient in writing to demonstrate their knowledge, plays a significant role in teaching writing. In academic settings, the importance of writing skills is magnified. Academic writing not only serves as a means of communication but also as a tool for critical thinking and knowledge dissemination. Bailey (2003) states that academic writing, tailored for the ones studying English-medium institutions, requires students to write essays and assignments with the complex knowledge of academic vocabulary, referencing and organisation. With all these requirements, teaching academic writing is highly demanding for both learners and teachers, as well. Kroll (2001) states that the responsibility of teachers is substantial since they provide the instruction to address these diverse needs for learners. Thus, the demands of academic writing underscore the critical role of both learners and

teachers in fostering the literacy skills essential for success in academic and professional contexts.

2.1.1. Approaches to Teaching Writing

In the realm of teaching academic writing, three predominant approaches have emerged: product writing, process writing, and the genre approach followed by genre process approach. Nunan (2015) introduced two approaches including product and process writing. While the product and process approaches have predominantly influenced EFL (English as a Foreign Language) and ESL (English as a Second Language) classrooms over the past two decades, the genre approach has gained considerable attention and importance in the last 10 years (Hyland, 2002).

Product Writing

In the product approach, also known as the traditional approach, learners are required to imitate previously presented texts (Gabrielatos, 2002). In this approach, learners are expected to adhere to predetermined rules regarding grammar and style (Brown, 2000). Nunan (1999) explains that students should create perfect copies of the provided writings under these principles. Badger and White (2000) also adopted the product approach to illustrate the learning stages of writing, which include familiarization, controlled writing, guided writing, and free writing. They describe a typical product class with the example of describing a house: starting with vocabulary related to houses, forming short sentences about the house, writing a text based on a picture of a house, and finally describing their own house. In summary, the product writing approach focuses on the comprehension of language structures with an emphasis on mastery of structural knowledge. However, the product approach has been criticized for its sole focus on text while excluding the process (Cribb, 2002).

Process Writing

Adding more on this criticism, Matsuda (2003) stated that instead of regarding writing as a mere reproduction of studied syntactic or discourse structures, the process

writing regards writing as a developmental process focused on organization of ideas and meaning. Kroll (2001) defines process approach as a cyclical process rather than a one-time effort including the stages of drafting, receiving feedback from peers or teachers, and revision rather than producing and submitting studied texts. Similar to Kroll's definition, Steele (2004) includes eight stages for process writing including brainstorming the possible ideas, planning, and structuring the ideas emerged during the brainstorming, mind mapping for connections, creation of the first draft, receiving feedback, editing, and submitting the final draft with the evaluation and teachers' feedback. Regarding these stages included in the process writing, it is possible to state that process approach is related with cognitive processes (Grabe and Kaplan, 1996). Furthermore, the process approach acknowledges that writing is not a solitary activity (Badger and White, 2000) but often involves interaction with others (Steele, 2004). Through peer feedback and teacher guidance, students can gain new perspectives on their work, identify areas for improvement, and refine their writing skills (Myles, 2002). However, process approach was also criticized because of its view of regarding writing as a one-piece entity which was regarded the same for writing academic essay and writing a post-card; however, the two types of such writings are not the same regarding the processes adopted (Badger and White, 2000). Furthermore, Badger and White (2000) noted that social aspect of writing was excluded in process approach. Graham and Harris (1997) also add that process writing did not prepare students for the difficulties of writing since it had the monolithic understanding of writing. In conclusion, while the process approach promotes developmental and interactive writing, it is criticized for its monolithic view and lack of consideration for diverse contexts and social aspects of writing.

Genre Approach

In the light of these criticisms, the genre approach has emerged, which could be accepted as an addition of product writing approach according to Badger and White (2000). Richards and Schmidt (2010) define genre as a form of discourse taking place in a specific context with its unique and identifiable designs and organizations, structures, and communicative purposes. Paltridge (2004) highlights that the teaching of these specific context, mentioned genres, is crucial for students' following communicative achievement. Hyland (2003) asserts that the genre approach is rooted in the social context where texts are

created and utilized; as a result, texts develop specific structures and rules shaped by the society, facilitating connections between similar texts. Badger and White (2000, p. 156) provide an example of a genre-based writing class where students study real estate descriptions, focusing on language patterns and social context. Initially, they create partial descriptions with support, then independently write full descriptions that mirror the original texts' context and style.

Adding more on genre approach, three approaches were evidently identified (Hyon, 1996). The first one, the New Rhetoric approach, focuses on the contexts of genres' rhetoric rather than the elaborate analysis of the elements of the texts (Freedman and Medway, 1994) to connect the type of the text and the condition of rhetoric (Coe, 2002). The second one, ESP (English for Specific Purposes), focuses on linguistic features, evaluating genre as communicative acts for communities sharing broad social aims (Swales, 1990). Thus, according to ESP theorists, to teach and evaluate the academic and professional language of nonnative speakers, ESP is a beneficial approach (Swales, 1990). The third one, Systemic Functional Linguistics (SFL), focuses on the context with the help of lexio-grammatical and features of rhetoric (Martin, 1997). Martin states that SFL is "concerned with showing how the organization of language is related to its use. In SFL this concern is pursued by modelling both language and social context as semiotic systems in a relationship of realization with one another" (p. 3). However, although genre approach, with all of its sub-approaches under, had the advantage of contextualising the text for both readers and purpose, it might demand for rhetorical comprehension of texts causing prescriptive instruction of them (Hyland, 2003). In conclusion, although the genre approach effectively contextualizes texts, it may require extensive rhetorical comprehension, potentially leading to prescriptive instruction.

Genre Process Approach

After all of these critiques, new approach named as the process-genre approach emerged. Uzun (2024) states that the process-genre approach acknowledges the importance of syntax from the product approach while also emphasising the mental processes involved in writing by integrating the genre approach while considering the social context and conventions of different discourse communities by combining the strengths of previous

methods rather than rejecting them. As stated by Badger and White (2000) in the genre approach with an example of real estate, in the process genre approach, teachers should closely mimic real-life situations and support learners in understanding the purpose and social context of their writing. For example, aspiring estate agents need to craft house descriptions aimed at selling (purpose) to specific audiences (tenor), include relevant details (field), and follow conventional formats (mode). Using their knowledge of vocabulary, grammar, and organization, learners can then draft and revise their descriptions to accurately reflect the real-world context (Badger and White, 2000). Engaging in such processes can enhance the understanding of various text types and the overall writing process (Guo, 2005). Badger and White (2000) summarise the stages of writing in process genre approach as modelling, construction with guidance and independent construction similar to the implementation in the classroom including replication of real situations, provision of the input from the teacher, and facilitation for skills; thus, enabling holistic learning of writing skills, linguistic and contextual aspects of different genres. In conclusion, the Process Genre Approach provides a comprehensive framework for teaching writing by combining the strengths of product, process, and genre approaches. It emphasizes the importance of context, purpose, and linguistic features while also focusing on the stages of writing and developing writing skills through guided practice and independent construction of texts.

In summary, the evolution of teaching academic writing has transitioned from the product approach, which emphasizes imitation and structural mastery (Brown, 2000; Gabrielatos, 2002), to the process approach, which focuses on the developmental and interactive aspects of writing (Grabe and Kaplan, 1996; Kroll, 2001). The genre approach further contextualizes writing by considering the social context and communicative purposes of different genres (Hyland, 2003; Paltridge, 2004). Recognizing the strengths and limitations of these approaches, the process-genre approach has emerged as a holistic method, integrating the structural focus of the product approach, the developmental processes of the process approach, and the contextual understanding of the genre approach (Badger and White, 2000; Uzun, 2024). This comprehensive framework not only enhances students' writing skills but also prepares them for the diverse and dynamic demands of real-world writing.

2.1.2. Sources of Feedback

The various approaches to teaching writing, including the product, process, and genre approaches, were highlighted so far with the importance of structured guidance and practice in developing students' writing skills. Feedback serves as an essential tool (Hyland and Hyland, 2006) in the teaching of academic writing (Morton et al., 2015), providing learners with the insights and directions to enhance linguistic accuracy (Ene and Upton, 2018) while the students had the chance to realise and fix their errors (Bitchener and Knoch, 2010) . Within the sources of feedback, both teacher feedback and peer feedback were examined.

Teacher Feedback

Teacher feedback, encompassing written corrective feedback (WCF), written commentary (WC), and oral feedback (OF), supports learners by improving their writing performance (Hyland and Hyland, 2019; Razi, 2023; Zhang, 2013). To begin with, Florio (2023) distinguishes key principles for effective teacher feedback, stating that it should be clearly stated, specific, and timely—neither too late to avoid forgetting the task nor too early to ensure the session is complete. Similarly, Hattie (2012) emphasizes that teacher feedback should be positive to make learners feel accepted, model effort, motivation, and commitment for demotivated learners, and be respectful and patient with errors.

Having established these foundational principles of teacher feedback, it is crucial to delve deeper into each type of teacher feedback. Teacher-provided written corrective feedback (WCF) plays a crucial role in enhancing students' writing skills by offering insights into their work's structure and content while promoting linguistic accuracy and overall communicative competence. WCF carries substantial information about the structure and content of learners' work, aiming to help students improve their writing (Hyland and Hyland, 2019) by benefiting from Vygotsky's (1978) zone of proximal development (ZPD) theory. Additionally, Hyland and Hyland (2019) highlight that teacher feedback is more than just a score; it prepares students for future writing tasks and helps them understand how their writing is perceived by others.

Moreover, Lee (2020) defines WCF as error feedback focused on detecting micro language errors and emphasizes that students expect such feedback from their teachers. The absence of this feedback can negatively impact student motivation. Moser (2020) describes WCF as a quick method for indicating and responding to learners' errors, sometimes including corrections and sometimes not. However, for WCF to be effective, it is crucial that students produce a second draft reflecting on the feedback; otherwise, the feedback is often ignored (Irwin, 2017). This iterative process is essential for enhancing students' writing skills and ensuring that they internalize the corrections provided (Hyland and Hyland, 2019; Moser, 2020).

Furthermore, Moser (2020) suggests that WCF is best utilized with more advanced learners who understand the rules but cannot apply them autonomously. Nevertheless, focusing too much on written errors limits the time teachers can spend addressing other issues such as content and organization. In conclusion, effectively integrating WCF into teaching practices requires a balanced approach that addresses both language accuracy and broader communicative competence, ensuring that students not only correct their errors but also develop holistically as writers.

Teacher-written commentary serves as a pivotal educational tool since it fosters constructive dialogue and aids in both academic development and interpersonal relationships within the classroom. The utilization of written commentary by teachers aims to assist students in revising and enhancing their work through comments that are intended to be read, understood, and acted upon as described by Lee (2020). Consequently, this form of feedback transforms into a dynamic exchange between teachers and students since the commentary is influenced not only by the text but also by an understanding of each student's personality and overall classroom experience, as stated by Hyland and Hyland (2019).

Additionally, the focus of written commentary further categorized into areas such as content, organization, language, and genre issues by Lee (2020) and it also includes various forms including statements, questions, imperatives, and hedges as noted by Sugita (2006). Moreover, Moser (2020) emphasizes the value students place on personalized written

feedback, which is particularly a type of feedback that addresses them by name and outlines the strengths and weaknesses of their work. This is the vice versa of the feedback that is vague and generic, which tends to be undervalued (Weaver, 2006).

The clarity and understandability of commentary are critical since they must be sufficient for students to act upon (Moser, 2020). Lee (2020), Goldstein et al. (2006), and Hyland and Hyland (2006) highlighted that the effectiveness of written commentary relies on its specificity, concreteness, details, text-specificity, and practicality, avoidance of ambiguities and overly broad generalizations. However, it is also crucial for teachers to maintain a balance between criticism and praise since excessive criticism can lead to confusion and demotivation, rather than fostering an encouraging environment (Lee, 2017). Thus, effective commentary should not only provide praise, criticism, or suggestions but it should also engage students and cultivate relationships, thereby it could enrich the educational experience (Goldstein et al., 2006).

Finally, Hyland and Hyland (2019) note that such commentary goes beyond simple evaluative remarks with the aim of involving nuanced social interactions that significantly influence both the impact of the feedback and the overall teaching efficacy. In conclusion, written commentary serves as a pivotal educational tool and it fosters a constructive dialogue between teachers and students that not only aids in academic development but also enhances interpersonal relationships within the classroom (Goldstein et al., 2006; Hyland and Hyland, 2019; Lee, 2017).

Furthermore, oral feedback, despite its logistical challenges, plays a crucial role in the feedback process by offering personalized and interactive learning opportunities that can significantly enhance students' writing skills. Ferris (2014) asserts that written feedback is most effective when supplemented with oral feedback. Similarly, Nassaji and Swain (2000) view oral feedback as a collaborative activity where both the learner and the teacher actively and significantly engage with each other. However, Lee (2017) highlights that the implementation of oral feedback faces several challenges due to time constraints and the large number of students in classes. Consequently, covering a set curriculum while also

providing timely written feedback on students' assignments places a high demand on teachers (Lee, 2017).

Moreover, during oral feedback sessions, students have the opportunity to reflect on the feedback they have received, discuss it further with their teacher, and ask for additional clarification or suggestions on how to apply the feedback to improve their writing (Lee, 2020). Teachers, in turn, can offer personalized tutoring, expand on their written feedback to deepen the student's understanding, and promote a positive learning environment (Lee, 2020). Specifically, Lee (2020) refers to oral feedback sessions as post-written conferences and outlines the features of effective oral feedback. These features include consistency, systematic organization, specific concentration on predetermined aspects, solution-focused advice, assistance to students, encouragement of active student involvement, and conduct in an environment that fosters rapport and mutual respect (Lee, 2020).

Nevertheless, Williams (2012) notes that oral feedback can become less accessible after the moment it is given. In addition, Oliver, and Adams (2021) discuss oral corrective feedback, which occurs when teachers or peers respond to language learners whose spoken output is incorrect, non-standard, inappropriate, or unclear. This type of feedback signals to learners, through spoken means and either implicitly or explicitly, that there is an issue with their utterance (Oliver and Adams, 2021). While it is relatively straightforward to deliver oral feedback in a classroom, ensuring it genuinely aids learners in enhancing their performance and successfully accomplishing the task is more challenging (Oliver and Adams, 2021).

Notably, Florio (2023) finds that oral feedback, particularly when posed as questions to individual students, has shown to be most effective. In conclusion, while oral feedback presents several logistical challenges, it remains a vital component of the feedback process, offering unique opportunities for personalized and interactive learning that can significantly enhance students' writing skills (Lee, 2020; Oliver and Adams, 2021).

In summary, effective teacher feedback—whether written corrective feedback, written commentary, or oral feedback—enhances students' writing skills and academic growth. Key principles include clarity, specificity, timeliness, positivity, and respect. While written feedback offers detailed corrections and insights, oral feedback provides personalized and interactive learning.

Peer Feedback

Also known as “peer response, peer review, peer revision, peer editing, peer tutoring, peer critiquing, and writing groups” (Hu, 2019, p. 97), peer feedback carries great importance for learners since Nuthall (2007) states that learners live in three different realms: the realm of the teacher, the impactful realm of peers, and the personal realm created by the personal experiences of students. Consequently, it is extensively utilized in writing classes (Storch, 2019). Peer feedback consists of providing comments and reviews on their peers' writing in the process of writing in oral, written, or online ways (Liu and Hansen, 2002).

To highlight the benefits of peer feedback, it is important to note that while there are numerous advantages for writing and learning of L2 (Ferris, 2003; Rollinson, 2005), learners have stated concerns about the usefulness and trustworthiness of peer feedback (Yoshida, 2008). Specifically, Lee (2020) summarizes the benefits of peer feedback as including different feedback focuses by teachers and peers (Li and Liu, 2010), complementing teacher feedback (Nelson and Schunn, 2009), enhancing audience awareness (Kaufman and Schunn, 2011), developing critical reading skills and self-critique (Lundstrom and Baker, 2009), promoting reflection (Patchan et al., 2009), and facilitating social interaction (Lu and Law, 2012).

However, it is crucial to address the challenges of peer feedback, which include a lack of feedback skills among peers (Li et al., 2012), a task-oriented mindset to please the teacher (Kaufmann and Schunn, 2011), lack of trust among students (Lu and Law, 2012), avoidance of negative comments (Liu et al., 2010), inhibition in L2 expression (Lundstrom and Baker, 2009), and superficial feedback (Patchan et al., 2009). Thus, while peer feedback

is a valuable tool, its effectiveness depends on addressing these challenges through structured training and support.

To strategize for effective peer feedback, Topping (1998) emphasizes the role of the process by which peers convey their ideas about the written products of their peers. At this point, Min (2005) explains a procedure for providing feedback to peers that consists of four stages: clarification of the writers' intentions, identification of problems, explanation of the problems, and suggestions to the problems. Similarly, Lee (2020) offers the praise-critique-explain-suggest rule, which includes a positive comment first, then identifying and stating the area needing improvement, describing the problem, and providing a practical suggestion for improvement.

Furthermore, to underscore the importance of peer feedback training, it is vital to note that, as Min (2016) asserts, peer feedback training is essential for successful implementation. Indeed, Lee (2020) identifies six tips integral to comprehensive peer feedback training. These tips include clarifying the purposes of peer feedback, demonstrating effective peer feedback behaviours, reflecting on received feedback to analyse and critique the feedback behaviours, providing explicit instructions on exemplary peer feedback behaviours, offering practice opportunities for peer feedback, and conducting feedback sessions from the teacher on the feedback from peers (Lee, 2020).

Furthermore, to enhance the process of peer feedback, Lee (2017) also suggests use of peer feedback sheets which are in the forms of open-ended comments on different aspects of peers' writing and the use of rating scales that require students to respond to selected aspects. Consequently, effective peer feedback could rely on structured procedures, comprehensive training, and well-designed feedback tools while ensuring the provision and reception of meaningful and constructive critiques.

In summary, peer feedback plays a vital role in writing classes, particularly for L2 learners since it fosters diverse perspectives and enhancing critical skills. Despite its benefits

such as improving audience awareness (Kaufman and Schunn, 2011), and promoting reflection (Patchan et al., 2009); however, concerns about its reliability persist (Yoshida, 2008). Therefore, effective peer feedback training that includes clear guidelines (Min, 2016), modelling, and practice (Lee, 2017) is essential for maximizing its benefits. The praise-critique-explain-suggest model (Lee, 2020) and the use of structured feedback sheets further support this process (Lee, 2017), making peer feedback a powerful tool for learning and development when properly implemented.

2.1.3. Ways of Giving Feedback

As the various sources of feedback were discussed including teacher and peer feedback, and their respective roles in enhancing students' writing skills, it is essential to examine the different ways in which this feedback can be delivered. The effectiveness of feedback is not only determined by its content but also by the method of delivery. Different methods offer unique advantages and challenges, and consequently, this impacts how students receive, process, and act upon the feedback. For this reason, this section explored the ways of giving feedback by focusing on face-to-face, online, synchronous, and asynchronous methods.

Face-to-face Feedback

As suggested by its name, face-to-face feedback includes the mutual dialogue between the feedback provider and the receiver and face-to-face feedback can be given by teachers, peers, or both. In the current literature, several aspects of face-to-face feedback from different providers have been reported by several researchers. For example, Kerr (2020) lists the benefits of face-to-face peer review as the enhancement of understanding and engagement with written feedback, responsibility for the feedback process, and the non-anonymity of face-to-face feedback. Adding on these benefits, Kerr (2020) states the advantage of face-to-face feedback as the enhancement of understanding and engagement by further elaborating on the written feedback, and; thus, making it a complementary process. Additionally, Kerr (2020) states the accountability of the face-to-face feedback by

stating that the non-anonymity of the peer requires the peer to provide feedback more seriously compared to an unidentified peer. As another advantage of face-to-face feedback, the ability to interpret nonverbal signals like body language and facial expressions enriches the interaction by providing immediate feedback and making the conversation more memorable (Meyer, 2006). It can also enhance the learning experience and help students better remember details and who said what (Meyer, 2006). Supporting Meyer's (2006) statements, Liu et al. (2020) assert that face-to-face feedback allows learners to explore issues in greater detail and more thoroughly thanks to mutual dialogue. As a positive effect of face-to-face feedback on other skills, DiGiovanni and Nagaswami (2001) also noted that face-to-face dialogue helps students improve their oral feedback skills.

However, there are notable weaknesses associated with face-to-face feedback. Liu et al. (2020) state that face-to-face feedback can be time-consuming because of the need for scheduling, physical arrangements, and the allocation of time of the instructor. Similarly, Crews and Butterfield (2014) identify several weaknesses of face-to-face feedback. For example, learners might avoid giving direct comments and provide superficial or overly positive feedback to maintain group harmony and their friendship, relationship with their peers. Additionally, some learners might be reluctant to receive criticism in person from other learners which might lead to negative feelings. Furthermore, characteristics such as shyness and avoidance of confrontation may result in less honest and critical feedback. Learners may also refrain from providing feedback to peers to avoid offending them (Crews and Butterfield, 2014). These weaknesses can be applicable to teacher face-to-face feedback as well. In conclusion, while face-to-face feedback offers several benefits such as enhanced understanding, immediate clarification, and improved oral feedback skills, it also presents challenges like time consumption, superficial feedback, and the potential for negative emotions.

Online Feedback

Online feedback, as defined by Narciss (2008), is information delivered through online platforms to inform learners about their current learning and performance status.

Specifically, in ESL/EFL writing, online feedback involves evaluations and comments on students' writing provided by teachers or peers with the aim of improving learners' writing skills (Hyland and Hyland, 2019), either synchronously or asynchronously.

To elucidate online feedback, also referred to as e-feedback, it is essential to comprehend both synchronous and asynchronous e-feedback. Synchronous e-feedback, according to DiGiovanni and Nagasawami (2001), enables learners to stay engaged with the task and receive real-time guidance from their instructor. Beauvois (1998, p. 198) metaphorically describes synchronous e-feedback as “a conversation in slow motion.” Synchronous e-feedback can thus be defined as an immediate response to the learners' written products. Several studies have explored the pros and cons of synchronous feedback. Advantages include metalinguistic corrections (Morris, 2005; Sauro, 2009) and improvements in vocabulary and grammar knowledge (Schultz, 2000). However, synchronous e-feedback may lead to an overemphasis on vocabulary and grammar (Blake and Zyzik, 2003) and may not be efficient in terms of time management and peer review (Schultz, 2000).

Conversely, asynchronous e-feedback, which includes methods such as email, discussion board messages, or comments/track changes in Microsoft Word (Ene and Upton, 2018, p. 2), can be defined as non-simultaneous or non-concurrent feedback. Advantages of asynchronous e-feedback include its effectiveness in improving micro skills such as grammar, spelling, and vocabulary (Vinagre and Munoz, 2015). However, disadvantages include its slower nature, lack of interaction, and insufficiency in encouraging in-depth revisions (Ho and Savignon, 2007).

In summary, online feedback enhances ESL/EFL writing skills through synchronous and asynchronous methods. Synchronous e-feedback offers immediate guidance, aiding real-time corrections but sometimes overemphasizing grammar. Asynchronous e-feedback improves micro skills with detailed responses but lacks interactive depth.

2.1.4. Types of Feedback

In this section, feedback was handled under two types including corrective and constructive feedback on writing.

The first feedback type is corrective feedback also called as error feedback. Corrective feedback can be categorized into two main types: explicit (direct) feedback, which directly corrects the learners' errors, and implicit (indirect) feedback, which incorporates the correct form into the response and may or may not include a metalinguistic explanation. Additionally, corrective feedback can be focused, targeting specific errors, or unfocused, addressing a wide range of errors.

Lightbown and Spada (2006) define corrective feedback under two aspects: first is explicit (direct) corrective feedback directly correcting the learners' error; second is implicit (indirect) corrective feedback incorporating the correct form into the response including or not including the metalinguistic explanation of the rule. Explicit feedback involves presenting the correct form of an error, which can include striking through, erasing, or adding the correct element (Bitchener et al., 2005). This type of feedback is particularly beneficial when students are unaware of the correct form (Ellis, 2009), making it more effective at lower proficiency levels (Bitchener, 2008). Additionally, it has been claimed that explicit feedback is effective in learning specific grammar points (Sheen, 2007). In contrast to explicit feedback, implicit feedback involves indicating that a student has made an error by underlining, circling, or marking it with an arrow and adding a note (Ferris and Roberts, 2001). This approach requires students to analyse and correct the error themselves. linguistic knowledge demands, implicit feedback is more commonly preferred at higher proficiency levels (Ferris, 2002). Another point in corrective feedback is the feedback including metalinguistic information, which was defined as giving clues about the type of error and how to correct it and aiming to engage learners' explicit knowledge, helping them understand the nature of their mistakes (Ellis et al., 2008). Explicit feedback is more effective at lower proficiency levels, while implicit feedback is preferred at higher levels, and metalinguistic information aids in understanding errors.

Another distinction on corrective feedback includes focused and unfocused feedback (Bitchener and Knoch, 2008). As explained by Ellis et al. (2008), unfocused feedback, also called as extensive feedback, includes the correction of wide range of error on the written product and is not recommended by L2 writing researchers. On the other hand, focused corrective feedback aims to correct specific errors like past simple tense errors rather than addressing wide range of errors like in the unfocused feedback (Ellis et al., 2008). In short, focused feedback targets specific errors, making it more effective than unfocused feedback, which corrects a wide range of errors.

Constructive feedback, defined as “sensitive and difficult art in teaching” by Kilbourne (1990, p. 344), is a fundamental element of the L2 writing classrooms for decades (Hyland and Hyland, 2019). Hyland and Hyland (2019) state that learners appreciate and expect constructive and heartening comments rather than cliches. However, this might be difficult for teachers because of the features of constructive feedback being both motivating and instructive (Fong et al., 2018). Thus, to minimize negative emotions, constructive feedback should avoid remarks that might be perceived as unhelpful or unnecessary by refraining from a patronizing tone towards the recipient (Fong et al., 2018). Hyland and Hyland (2019) supported this idea, and they emphasized the balance of positive and critical feedback that helps learners to build confident and informed writing skills with the help of “constructive and forward-looking feedback” (p. 3). With the statement of balanced feedback, Hyland, and Hyland (2019) elaborate that while negative feedback could harm writer’s confidence early and unnecessary praise might confuse students and discourage them from revising. In other words, the way feedback is given can build relationships between the teacher and the learners that either support or impede a learners’ progress in writing. In summary, constructive feedback, when delivered with a careful balance of positivity and critique, plays a pivotal role in fostering students’ confidence and skills in L2 writing, ultimately enhancing their academic progress and the teacher-student relationship.

Regarding the provision of the constructive feedback, Fong et al. (2018) provide recommendations for providing constructive feedback including five appraisal categories. These categories consist of providing clear paths to improvement, which helps students see specific ways to enhance their work; avoiding mismatches between feedback and task to

prevent confusion and frustration; focusing on the student's ability or self to support self-efficacy; highlighting the relationship between feedback giver and receiver to build trust and positive emotional responses; and valuing the task to enhance motivation and engagement (Fong et al., 2018). In conclusion, the recommendations by Fong et al. align well with effective practices in academic writing feedback discussed so far.

In summary, both corrective and constructive feedback are essential in enhancing students' writing skills. Corrective feedback, whether explicit or implicit, helps students identify and rectify their errors, with focused feedback proving more effective than unfocused feedback while constructive feedback, balancing positivity with critique, fosters confidence and motivation in learners.

2.2. Anonymous-Multi Mediated Writing Model

Taking into the various feedback methods discussed consideration, it is evident that innovative approaches are essential in fostering more effective writing skills. One of these innovative approaches is the Anonymous Multi-Mediated Writing Model (AMMW) developed by Razi (2017, 2023). The AMMW model leverages the advantages of anonymity and multiple feedback sources to improve the quality and reliability of peer reviews. This model addresses some common challenges associated with traditional peer feedback methods and it also provides a structured, balanced, and effective way to enhance academic writing skills. In short, the AMMW Model provides a novel approach on improving academic writing through anonymous and diverse feedback mechanisms. Razi received the Turnitin Global Innovation Award (Europe Student Engagement) in 2015 as the model promotes academic integrity and helps the prevention of plagiarism and published extensively on various implementations of the model at different settings. Following sections presents selected related publications from Razi on the implementation of AMMW Model.

2.2.1. Study by Razi in 2016

Razi (2016) examined the connection between review skills and academic writing skills by employing various feedback methods which includes self-review, peer review, tutor review, and anonymous digital feedback through text-matching software such as Turnitin. The study included 87 first-year students from the ELT department, and the data collection tool included TAWR, which was also utilized in this current study. The data collection procedure involved students who submitted drafts and received feedback from peers, tutors, and the digital tool. Then, the students were required to review their own and peers' papers using TAWR. After reviewing the feedback, the scores which were collected via TAWR were used to compare the scores from the lecturer to ensure reliability and correlation. In light of these procedures, Razi (2016) reported key findings consisting of four core topics which include the connection between writing and reviewing skills, the impact of writing success on self-review skills, the impact of writing on peer review skills, and the effectiveness of multiple feedback sources. In conclusion, Razi (2016) highlights the interrelation between writing and reviewing skills and he showed that multiple feedback sources enhance academic writing quality and reliability.

The connection between writing and reviewing skills emphasizes that higher-scoring students provide more reliable self and peer reviews, and there was a weak positive correlation between students' academic writing and their reviewing skills, indicating that good writers tend to be better reviewers (Razi 2016). The impact of writing success on self-review skills suggests that good students, selected according to their writing scores from the categories of good, acceptable, and poor, showed more consistency between their self-scores and the lecturer's scores, indicating better self-assessment abilities (Razi 2016). Similarly, better writers performed better in peer reviews, and the anonymous peer review process facilitated by Turnitin proved beneficial as it reduced biases related to familiarity and relationships among peers (Razi 2016). Combining self, peer, tutor, and digital reviews helped students develop better academic writing skills, and digital tools like Turnitin with features like Peer Mark were instrumental in managing the peer review process and providing effective feedback (Razi 2016). As a conclusion, Razi summarized that academic writing and review skills are interrelated, with stronger students excelling in both areas. He

highlighted that using anonymous peer review which was supported by digital tools and the incorporation of multiple feedback sources improves the reliability and comprehensiveness of student assessments. Razi also provides implications for educators and suggests that they should incorporate review skills and digital tools like Turnitin into writing courses to enhance feedback quality, to foster learner autonomy, and to improve both writing and critical assessment abilities. Overall, the study underlines the beneficial impact of integrating multiple feedback sources and digital tools to enhance writing and reviewing skills in academic settings.

2.2.2. Study by Razi in 2017

On his second research study regarding AMMW model, Razi (2017) compared open and anonymous peer feedback in digital environments by implementing the anonymous multi-mediated writing model. It was also the first study including the name of AMMW model. Razi (2017) conducted the study with a true experimental design and also involved 60 students from an academic writing course. Participants were divided into two groups including the experimental group with anonymous peer review and the control group with open peer review. After participants submitted their drafts, they received feedback, and revised accordingly to resubmit their work. The feedback was provided and received based on TAWR, and the participants were categorised as good, moderate, or weak writers. Key findings from the study included the advantageous nature of anonymity bringing encouragement of honest critique, improved revision quality, balanced benefits across proficiency levels, and metacognitive skills development, along with challenges and suggestions. In short, Razi (2017) demonstrated the effectiveness of anonymous peer feedback on improving critique quality, revision processes, and metacognitive skills among students.

Anonymous feedback in a digital peer review environment enhances the overall quality and honesty of critiques as students feel less concerned about damaging their relationships with peers, and this leads to more genuine and constructive feedback (Razi 2017). Anonymous feedback approach significantly improves revision quality with the

students making better corrections than open feedback (Razı 2017). Moreover, learners across all proficiency levels including good, moderate, and weak benefit from this balanced distribution of feedback. Although weaker students may provide less effective feedback, they could still improve themselves and contribute useful insights (Razı 2017). This process also fosters the development of metacognitive skills by encouraging students to critically evaluate their own and the peer's work, which enhances their critical thinking and decision-making abilities in writing. In turn, this enhancement ultimately promotes more autonomous and reflective learning (Razı 2017). In conclusion, the study reveals that anonymous feedback not only enhances critique quality but also supports balanced benefits across different proficiency levels and fosters metacognitive development.

Regarding the challenges, anonymous feedback can sometimes lead to overly harsh critiques, so it is essential to train students on how to provide constructive and supportive feedback (Razı 2017). Additionally, the peer review process increases students' workload; therefore, incorporating the quality of peer feedback into final assessments can motivate students to participate thoroughly (Razı 2017). The study also found digital platforms effective for anonymous feedback and suggested that future tools should include features to group students by proficiency, thereby saving teachers' time and ensuring balanced feedback distribution (Razı 2017). Addressing challenges, the study recommends training for constructive feedback and leveraging digital platforms to streamline and balance the peer review process.

2.2.3. Study by Razı in 2023

In the study by Razı (2023), the AMMW model was adapted for emergency remote teaching during the COVID-19 pandemic. In this study, Razı (2023) aimed to explore the application of the AMMW model in an emergency remote teaching setting to maintain academic integrity and assess its effectiveness in a digital environment. Razı (2023) employed a mixed-methods research design for this study and conducted this study in a writing skills course for first-year students in Türkiye. Quantitative data included the feedback evaluations from the instructor, scores from writing assignments which were

scored by TAWR, and FAWR (Fundamentals of Academic Writing Rubric). FAWR was a simplified version of TAWR. On the other hand, qualitative data were derived from reflection papers and interviews. In conclusion, Razi (2023) evaluates the AMMW Model's adaptation to emergency remote teaching and demonstrated its effectiveness in maintaining academic integrity and student performance in a digital setting.

Regarding the findings, Razi (2023) stated that when emergency remote teaching and face-to-face teaching scores were compared, there was no significant difference found in student performance and instances of academic misconduct. Furthermore, the AMMW Model facilitated effective feedback mechanisms in emergency remote teaching, including asynchronous peer and teacher feedback. Regarding the qualitative findings, despite technological challenges and a preference for face-to-face interactions, students found the online adaptations effective, particularly they appreciated the asynchronous components that allowed flexibility and self-paced learning. As a conclusion, Razi (2023) stated that the AMMW Model's adaptation to emergency remote teaching was successful in maintaining academic integrity and student performance. The model could serve as an example for similar remote teaching adaptations in other academic disciplines or institutions facing abrupt transitions to online education. Ultimately, the findings suggest that the AMMW Model effectively supports academic integrity and student performance in remote teaching, and also it offers a blueprint for future adaptations.

2.2.4. Summary of Razi's Selected Studies

When all three research studies are inspected in detail, four common points emerged evidently. Firstly, the importance of peer review and feedback in improving academic writing skills are emphasized in all three studies since they discussed how structured peer interactions can lead to better writing outcomes and deeper engagement with the material. Secondly, there is a significant focus on the use of digital tools to facilitate feedback and peer review processes. For example, tools like Turnitin are highlighted for their ability to provide similarity check and structured peer feedback mechanisms, which are seen as beneficial for both learners and educators. Thirdly, the studies explored issues related to

teaching and improving academic writing within the context of ELT since they investigated how various interventions, including peer reviews, digital tools, and specific teaching methodologies, can enhance the writing skills of EFL students who are non-native English speakers. The fourth recurring theme is the importance of fostering autonomy among learners by engaging them in peer review processes with the help of utilization of technology that encourages independent learning and critical thinking. Overall, these studies collectively emphasized the critical role of peer review, digital tools, and autonomous learning in enhancing academic writing skills, particularly for non-native English speakers.

2.2.5. Processes of Anonymous-Multi Mediated Writing Model

AMMW model processes include steps of process writing consisting of “prewriting, planning, drafting, and revision originating from the teacher, peer, and self-feedback before final submission on a digital platform” (Razı 2023, p. 2). Razı (2017) states that the primary motivation behind creating the AMMW Model was to encourage students to diligently complete their writing tasks while they are gaining insights from various feedback sources, which included evaluations from teachers, peers, and their own self-assessments throughout the composition process in order to guide learners to thoughtful consideration at every stage. Razı (2017, 2023) adds that this process enhances the creativity. Before submitting improved versions of their assignments, students incorporate feedback from their instructor on their initial drafts. After the revision and submission of the assignment, each assignment is matched with three peers, each from their writing proficiency categories of good, moderate, and weak. Each feedback was exchanged anonymously in case of being reluctant to criticize their friends. Following peer evaluations, students further refine and resubmit their work. Before the final submission of their assignments, students also conduct a self-assessment of their papers. Throughout this process, an assessment rubric serves as an educational tool to facilitate various types of feedback. In short, the AMMW model’s process involves multiple stages of feedback and revision, enhancing writing skills through a structured and anonymous peer review system.

2.2.6. Theoretical Background of Anonymous-Multi Mediated Writing Model

Although there are several theories supporting the model, three theories stand out as vital. As mentioned in Razi (2023), these include scaffolding, the ZPD, and metacognition. Regarding scaffolding, often linked to the Soviet psychologist Vygotsky, whose research significantly influenced the evaluation of scaffolding, Bruner, and Sherwood (1976) examined interactions of mothers and their children during peekaboo games and described the mothers' supportive actions as scaffolding. Following this study by Bruner and Sherwood (1976), Wood et al. (1976) applied the term in educational discourse with the supportive role of parents in guiding their children through problem-solving tasks. Thus, Wood et al. (1976) defined scaffolding as a helping hand of an adult "that enables a child or novice to solve a problem, carry out a task, or achieve a goal which would be beyond his unassisted efforts" (p. 90). In conclusion, the concept of scaffolding, developed through educational research, highlights the supportive role of more knowledgeable individuals in helping learners achieve their goals.

Regarding the ZPD, which is also at the centre of scaffolding, Vygotsky (1978) defined it as "the gap between the current developmental stage achieved through independent problem-solving and the potential developmental stage attainable with guidance from adults or collaboration with more skilled peers" (p. 86). As stated by Gönülal and Loewen (2018), scaffolding is temporary and adaptable support provided in the ZPD, which in turn refers to the strategies that educators employ within the borders of learners' ZPD to enhance learning and develop their knowledge and skills. These theories find a place in the AMMW Model with the integration of feedback, which enables interaction with more knowledgeable peers (Razi 2023). Shortly, the ZPD emphasizes the gap between independent problem-solving and potential development, supported by scaffolding from more knowledgeable peers.

The theory of metacognition, which emerged in the mid-1970s by John H. Flavell, has a vital place in the AMMW model as stated by Razi (2023). According to Flavell (1979), "metacognitive knowledge consists primarily of knowledge or beliefs about what factors or variables act and interact in what ways to affect the course and outcome of cognitive

enterprises” (p. 907). On the other hand, the experience originating from metacognition consists of intellectual, conscious cognitive or affective experiences reflecting on one's cognitive processes like understanding or problem-solving.

In the AMMW Model, learners who receive feedback from peers are required to conduct analyses since they are responsible for managing the feedback application; thus, they are involved in metacognitive experiences (Razı 2016, 2017, 2023). In short, metacognition involves understanding and managing one's cognitive processes, and in the AMMW Model, this is facilitated through feedback and self-assessment, promoting deeper learning and reflection.

In conclusion, Razı's AMMW Model enhances academic writing by leveraging anonymous, multi-sourced feedback. It addresses traditional peer feedback limitations and improves review quality and reliability. Razı's selected studies from 2016, 2017, and 2023 highlight the model's effectiveness in interlinking writing and reviewing skills, benefiting from anonymous digital feedback, and adapting to remote teaching. Supported by theories like scaffolding, the ZPD, and metacognition, the AMMW Model fosters better academic writing and learner autonomy, offering a valuable tool for educators and students.

2.3. Generative Artificial Intelligence

Although it gained widespread recognition with the advent of ChatGPT, artificial intelligence (AI) and GenAI have been a significant presence in the world for many decades. The latest publication in The Economist Magazine (“A Short History of AI”, 2024) outlines the history of artificial intelligence and notes a pivotal meeting that occurred 69 years ago where the term ‘artificial intelligence’ was first mentioned. Subsequent to this event, the literature began to highlight the achievements of Turing Award winners. Thus, the foundations of AI were laid long before the emergence of contemporary AI models like ChatGPT.

Marvin Minsky (1969) and John McCarthy (1971) for defining the foundations of the field based on representation and reasoning; Allen Newell and Herbert Simon (1975) for symbolic models of problem solving and human cognition; Ed Feigenbaum and Raj Reddy (1994) for developing expert systems that encode human knowledge to solve real-world problems; Judea Pearl (2011) for developing probabilistic reasoning techniques that deal with uncertainty in a principled manner; and finally Yoshua Bengio, Geoffrey Hinton, and Yann LeCun (2019) for making 'deep learning' (multilayer neural networks) a critical part of modern computing. (Russell and Norvig, 2021, p. 35)

In recent years, a significant development emerged, leveraging all the advancements discussed by Russell and Norvig (2021). This development, met with considerable interest, was the introduction of the Generative Pre-trained Transformer 2 (GPT-2) by OpenAI in February 2019. Trained on an extensive dataset of 1.5 million samples, GPT-2 demonstrated impressive capabilities across various language tasks. According to IBM ("What are LLMs?", n.d.), large language models (LLMs) like GPT-2 are foundational models trained on vast datasets, enabling them to comprehend and generate natural language as well as other forms of content. Following GPT-2, OpenAI launched GPT-3 in June 2020 with the feature of 175 billion parameters and enhanced language creation capabilities. The subsequent release of GPT-4 and later models in 2023 marked a new era since new GPT models could search for sources available on the internet and gather information to respond to given prompts. Following OpenAI's pioneering efforts with ChatGPT, other companies also made significant strides in this field such as Google with Gemini, Microsoft with Bing, and Anthropic with Claude, and numerous other AI tools available today. The competitive landscape of AI development has expanded significantly with numerous companies innovating and contributing to the field and increasing every day.

In conclusion history of AI and GenAI span decades starting from foundational contributions from Turing Award winners shaping the field. OpenAI's GPT-2 and GPT-3 significantly advanced AI language processing using vast datasets. The rapid and evolving nature of GPT models along with innovations from well-known companies like Google,

Microsoft, and Anthropic emphasises the competitive and transformative impact of AI technology.

2.3.1. Prompt Engineering

Built on the advancements in AI, the concept of prompt engineering has emerged as a crucial technique for effectively exploring the capabilities of large language models (LLMs). To understand prompt engineering, it is essential first to comprehend the meaning of a prompt which was defined as directives provided to LLMs to apply guidelines, manage tasks, and ensure certain characteristics and quantities of the generated output according to White et al. (2023). In short, they serve as a form of programming that can tailor the LLMs' responses and interactions (White et al., 2023). This foundational understanding underscores the significance of prompts in harnessing the full potential of LLMs.

Giray (2023) uses the analogy of a library to define prompt engineering. This analogy resembles directly instructing a language model to asking a librarian for a book, and the prompt engineering is likened to placing specific questions on bookshelves to guide the librarian to find the information. Ye et al. (2023) describe prompt engineering as a demanding but essential concept for enhancing the performance of LLMs. Prompt engineering involves sophisticated reasoning to analyse the LLMs' mistakes and to identify the lacking or misleading elements in the prompt, and to articulate or rearticulate the prompt (Ye et al., 2023). Similarly, Gu et al. (2023) succinctly define prompt engineering as an approach to offer guidance and input to a pre-trained model to apply its current knowledge to solve a new task. These definitions collectively emphasize the critical role of prompt engineering in refining and improving LLM performance.

The advantages of prompt engineering are also highlighted under several points. For instance, Giray (2023) notes that prompt engineering enables writers and researchers in academic fields to fully benefit from the capabilities of language models across diverse disciplines. Furthermore, Gu et al. (2023) state that with the help of prompt engineering, a minimal amount of labelled data is needed for a pre-trained model to work on new tasks,

which saves time and computing power. Additionally, Chen et al. (2024) assert that prompt engineering allows the full potential of LLMs to be achieved and applied in various fields. These benefits underscore the significant impact of prompt engineering on the efficiency and effectiveness of LLM applications.

In summary, prompt engineering has emerged as a vital technique for maximizing the potential of large language models (LLMs). Prompts, as directives that guide LLMs, are essential for tailoring responses and managing tasks. Analogies such as Giray's (2023) library comparison illustrate how prompt engineering goes beyond direct instructions by strategically guiding LLMs to find and generate relevant information. This process, as described by Ye et al. (2023), involves sophisticated reasoning to refine prompts and enhance model performance. As advantages, prompt engineering enables comprehensive utilization of LLMs in academic and diverse fields, reduces the need for extensive labelled data, and ultimately unlocks the full capabilities of these advanced models. Collectively, these insights highlight the transformative role of prompt engineering in advancing AI applications.

2.3.2. Generative Artificial Intelligence Research in Education

As the rich history of AI and the critical role of prompt engineering in optimizing the performance of large language models were discovered, it becomes evident that these advancements are not just confined to the realms of technology and industry. Their impact extends to various domains including education. Bozkurt (2023a) highlights that the potential of AI to process language naturally has a transformative effect on education in both positive and challenging ways. As a positive aspect, technologies like AI offer educational benefits such as personalized learning and enhanced accessibility (Bozkurt, 2023a). Similarly, Escotet (2023) provides an optimistic view of AI's potential impact on higher education, and emphasizes its ability to provide personalized learning, improve educational planning and administration, and enhance distance learning through customized coursework and flexible scheduling. However, they also present challenges, including biases against AI, privacy concerns, and overreliance on technology (Bozkurt, 2023a). On the same vein, Escotet (2023) also stresses the need for ethical engagement and cautious application of AI.

Regarding educators' perspectives on GenAI integration, Chiu (2023) and Van Wyk (2024) conducted similar studies with educators. Chiu (2023) conducted research with 88 schoolteachers from various fields in Hong Kong to explore their perspectives on GenAI. The study found that educators see the integration of GenAI tools like ChatGPT and Midjourney into school curricula as a way to enhance AI literacy alongside traditional subjects since these tools personalize learning experiences, support diverse educational needs, and transform assessment methods and administrative tasks by automating routine operations. However, significant challenges include concerns over the reliability of AI-generated content, ethical issues, and the potential for students to become overly reliant on AI for assignments. Chiu advocates for updated educational policies to include AI literacy and more professional development for teachers in AI integration. Similarly, Van Wyk (2024) interviewed eight academics from the faculty of education across the African continent about their views on GenAI tools. The study revealed that while academics recognize the benefits of GenAI in education, they are concerned about issues such as academic dishonesty. To address these concerns, Van Wyk highlights the importance of establishing ethical guidelines and institutional policies. For example, academics are employing strategies to integrate these tools responsibly, including designing authentic assessments that are challenging for AI to replicate and promoting digital literacy to navigate the ethical use of GenAI. The study calls for ongoing research to further understand the implications of GenAI in educational settings and policy revisions to manage these technologies effectively.

Regarding institutional guidelines and policies, McDonald et al. (2024) focused on 116 higher education institutions' guidelines in the United States to understand how these institutions guide and regulate the use of GenAI tools like ChatGPT in teaching, learning, and research activities. The results show that around 63% of these institutions encourage its use by providing structured integration guidelines, sample syllabi, and curriculum activities. More than half address ethical concerns by focusing on diversity, equity, inclusion, and privacy. Responses vary from strong resistance due to fears of academic dishonesty to enthusiastic support for GenAI as a transformative educational tool. Universities are adapting their pedagogical strategies to include GenAI, modifying assessment methods, and incorporating new teaching tools to enhance learning outcomes. However, significant

concerns about the reliability of GenAI detection tools and their impact on academic integrity remain.

From unique perspectives, Michel-Villarreal et al. (2023), Bharadwaj et al. (2023), and Shaw et al. (2023) conducted research studies on GenAI and education. Michel-Villarreal et al. (2023) conducted a unique study by interviewing ChatGPT to gain insights into its functionalities and implications for higher education. The analysis highlighted several opportunities and challenges, which includes continuous accessibility, personalized tutoring, supplemental learning resources, enhancement of language skills, support for educators, fostering innovative learning experiences, and assistance in research activities. However, challenges include maintaining academic integrity, ensuring quality control, capturing nuanced educational needs, providing domain-specific expertise, and facilitating effective communication and collaboration. On the other hand, Bharadwaj et al. (2023) conducted a large-scale survey study with 2,048 students, 1,748 instructors, and 306 higher education administrators to identify the differences between student and institutional stakeholder experiences and preferences regarding digital learning in higher education. The study found that the use of GenAI tools like ChatGPT is increasing among students with many using these tools for various academic purposes despite concerns about academic integrity. This finding underscores the growing acceptance and integration of GenAI in educational environments and highlights the need for balanced approaches to address the associated challenges. Building on these foundational insights, Shaw et al. (2023) conducted a follow-up study to understand the longitudinal adoption and impacts of GenAI tools on teaching and learning within higher education environments. Collecting responses from 1,000 higher education faculty and 1,600 postsecondary students, Shaw found a significant increase in GenAI tool usage among students and faculty. Students view it positively for learning outcomes while faculty opinions have become more neutral or positive but both groups see proficiency in GenAI as crucial for future professional success. However, ongoing concerns about academic integrity related to the use of GenAI in writing assignments influence policymaking at both institutional and course levels. Students use GenAI for a wide range of academic and non-academic purposes, while faculty primarily use these tools to understand student interactions and to teach effective usage.

From students' and cultural perspectives, Chan, and Hu (2023) and Yusuf et al. (2024) conducted research studies with students regarding GenAI and education. Chan and Hu (2023) conducted a study with 399 undergraduate and postgraduate students in Hong Kong to explore their perceptions of generative AI technologies in higher education. The findings revealed that students are familiar with GenAI technologies and aware of their limitations and potential inaccuracies. They showed a positive attitude towards using GenAI in their education and future careers and appreciate its potential to provide insights and personalized feedback. However, they expressed concerns about over-reliance on GenAI, its impact on the value of university education, issues related to accuracy, transparency, privacy, ethics, and potential negative effects on critical thinking, creativity, and career prospects. From a multicultural perspective, Yusuf et al. (2024) conducted a study with 1,240 students and lecturers from higher education institutions across 76 countries to examine the usage, benefits, and concerns of GenAI in higher education. Participants showed high awareness and familiarity with GenAI tools, with a significant portion having prior experience and intention to continue using these tools for information retrieval and text paraphrasing. They stated that GenAI significantly enhances learning processes but also expressed concerns about academic dishonesty and the need for ethical guidelines. Benefiting from Hofstede Insights, Yusuf et al. (2024) found that cultures with high uncertainty avoidance are more likely to see GenAI as a threat to academic integrity and support strict regulations to mitigate these risks. Cultures embracing long-term commitments and respect for tradition value traditional academic practices but also recognize the potential benefits of GenAI since they have a balanced view. Cultures with low power distance support regulatory frameworks to ensure fairness in GenAI usage. Collectivist cultures favour regulatory policies to maintain group harmony and fairness, and high masculinity cultures are more open to GenAI innovations and supportive of ethical guidelines for its use.

Lastly, as a summary of the research trends, Bozkurt (2023b) conducted a systematic literature review with 513 research studies indexed in Scopus and published between 2022 and 2024 to understand the trends and patterns of generative AI research in educational praxis. The review revealed thematic structures in titles, identifying key and emerging themes from abstracts. Bozkurt (2023b) categorized the findings into seven main themes:

1. “Interaction and communication with generative AI-powered chatbots” (p. 264): Enhancing learning experiences, promoting engagement, and providing personalized support to students.
2. “Impact of LLMs and generative AI on teaching and learning” (p. 264): Transforming teaching methodologies, improving learning outcomes, and providing innovative educational tools.
3. “Conversational educational agents and their opportunities, challenges, and implications” (p. 264): Presenting opportunities and challenges in the educational context.
4. “Leveraging generative AI for enhancing social and cognitive learning processes” (p. 265): Enhancing both social and cognitive aspects of learning.
5. “Promoting AI literacy for unleashing future opportunities” (p. 265): Incorporating AI literacy into the curriculum to prepare learners for future advancements.
6. “Harnessing generative AI to expand academic capabilities” (p. 265): Using generative AI to enhance academic performance and support educational needs.
7. “Augmenting educational experiences through human-AI interaction” (p. 266): Recognizing the social nature of education and the promise of AI in augmenting educational experiences.

In summary, the integration of AI, particularly generative AI, into education offers transformative potential alongside significant challenges. AI’s ability to personalize learning, enhance accessibility, and support educational planning is widely acknowledged by researchers like Bozkurt (2023a) and Escotet (2023). However, concerns about biases, privacy, and overreliance on technology persist. Educators and institutions, as highlighted by Chiu (2023) and Van Wyk (2024), recognize both the benefits and ethical dilemmas posed by GenAI tools, advocating for updated policies and professional development. Research by McDonald et al. (2024) underscores the varied institutional responses, emphasizing the need for structured guidelines and ethical considerations. Studies focusing on student perspectives, such as those by Chan and Hu (2023) and Yusuf et al. (2024), reveal a positive outlook towards GenAI but also caution against potential drawbacks.

2.3.4. Ethical Considerations for the Use of Generative Artificial Intelligence in writing

Risko and Gilbert (2016) define cognitive offloading as using external tools to reduce cognitive load. On this point, Khan (2023) emphasizes that the use of GenAI tools in writing assignments can be seen as a form of cognitive offloading. However, this raises issues of acknowledgment, as students may not fully credit the AI tools used in their work. To address this, transparency in the use of GenAI tools and clear guidelines and policies are crucial (Khan, 2023). In 2014, long before the emergence of ChatGPT, Khan (2014) coined the term ‘e-cheating,’ defined as “the use of various forms of ICT to perform academic misconduct or dishonesty, either in or out of classrooms, to gain an unfair advantage” (p. 45). Supporting this definition by Khan (2023), the Committee on Publication Ethics (COPE, 2023) has taken a firm stance on the use of AI tools in academic writing. COPE (2023) emphasized that AI tools cannot be considered as authors because their reliability and accountability are not on par with human authors. Consequently, any use of AI tools must be transparently stated to maintain the integrity of academic work according to COPE (2023). COPE (2023) also highlighted the importance of transparency and clear attribution when AI tools are used in the research and writing process, ensuring that the contributions from these tools are properly acknowledged without misleading readers about the human authorship of the work. These concepts are particularly relevant in the context of GenAI tools since it underscores the importance of addressing potential academic misconduct facilitated by technology.

Regarding emphasizing the importance of ethical use of GenAI in the process of writing, several studies have conducted. First of these is a study conducted by Barrett and Pack (2023) involving 158 students and 68 ESL and EFL teachers in the United States to examine their perspectives on the use of GenAI tools in the writing process. The study found that both students and teachers generally agreed on the appropriateness of using GenAI for brainstorming ideas or as a model for various writing tasks. However, there was a notable difference in opinion when it came to using GenAI without disclosure. For example, students were more open to this idea than teachers. Additionally, although teachers were more into accepting of using GenAI to model feedback or for evaluation purposes, a significant number of them reported that their institutions lacked clear policies regarding AI use. This

underscores the pressing need for well-defined guidelines and professional development on the use of GenAI in educational settings.

With a similar methodology on collecting reflections of the participants, McGuire et al. (2024) developed a model for utilizing GenAI tools for feedback in higher education. The study was based on student reflections and ChatGPT-generated feedback samples in Michigan. Their analysis found that using ChatGPT for peer-simulated feedback offered immediate, detailed, and formative guidance. This approach blended creative freedom with valuable advice, and it helped students explore various research angles and maintain consistent workflow. Nevertheless, the study also highlighted significant ethical challenges, including the generation of inaccurate information (hallucinations), depersonalized learning experiences, reproduction of biases, and risks of plagiarism or cheating. Therefore, McGuire et al. (2024) emphasized the importance of developing students' abilities to critically evaluate AI-generated content, to recognize and to correct inaccuracies, and to understand ethical implications. They also noted the crucial role of instructors in fostering discussions on the ethical use of AI and creating assignments that require students to reflect on their use of AI tools.

As a conceptual discussion, Kumar et al. (2023) addressed key concerns related to intellectual property and academic integrity with AI-generated text. Their study highlighted the increased challenges of detecting plagiarism and contract cheating in the age of AI. To decrease these concerns, Kumar et al. (2023) recommended new assessment and evaluation strategies to ensure the originality of student work. They suggested that educational policies should focus on promoting ethical AI usage since AI tools can enhance or automate the writing process. As a result, there is a need to rethink traditional concepts of plagiarism and contract cheating. Furthermore, Kumar et al. (2023) advocated for educating students on academic integrity, and offer diverse assessment options to demonstrate learning, and design assessments that minimize dishonest behaviour.

In conclusion, the integration of GenAI tools in writing assignments raises significant ethical issues, particularly regarding transparency and acknowledgment of AI contributions.

Studies by Barrett and Pack (2023) and McGuire et al. (2024) highlight the need for clear guidelines and institutional policies, as well as the importance of addressing ethical challenges such as plagiarism, inaccurate information, and depersonalized learning. Kumar et al. (2023) stress the necessity of innovative assessment strategies and education on academic integrity to mitigate the risks of AI misuse.

2.3.5. Generative Artificial Intelligence Based Writing

The integration of AI in academic writing has also gained significant attention, with studies comparing automated writing evaluation (AWE) systems, generative AI, and AI-based writing assistants. The research studies by Godwin-Jones (2022), Escalante et al. (2023), Gayed et al. (2022), Alharbi (2023), Nguyen et al. (2024), and Taiye et al. (2024) highlight both the benefits and challenges of these technologies.

Godwin-Jones (2022) reviewed multiple research studies to evaluate the impact of automated writing evaluation (AWE) systems on second language writing instruction and student engagement. His findings indicate that while AWE systems are effective in providing accurate and timely feedback on grammatical and lexical errors, they fall short in addressing higher-order writing skills such as organization and argumentation. Moreover, student engagement with AWE systems is influenced by their trust in the technology and their ability to utilize the feedback effectively, and higher proficiency students benefiting more. Therefore, the role of teachers is crucial in supplementing AWE feedback to help students improve their writing. Nonetheless, challenges such as over-reliance on automated feedback, neglect of higher-order writing skills, and difficulty catering to multilingual and diverse learners were also highlighted in the study.

Furthermore, Alharbi (2023) analysed 104 peer-reviewed scholarly publications to evaluate AI-powered writing assistance tools in the classroom and categorized these tools into four main types including automated writing evaluation tools, automated writing corrective feedback tools, AI-powered machine translators, and GenAI tools capable of producing human-like texts. The study found that AI-powered tools significantly improve

the quality of students' writing by providing detailed and consistent feedback. However, the effectiveness of these tools varies depending on students' proficiency levels and their ability to understand and use the feedback.

In addition, Nguyen et al. (2024) explored human-AI collaboration in academic writing, focusing on how doctoral students use generative AI tools like ChatGPT and its impact on writing performance. The study involved ten doctoral students from Finland and New Zealand who were tasked with writing essays on AI in education by using ChatGPT. The findings revealed two main strategies including structured adaptivity and unstructured streamline. Structured adaptivity, where students engaged deeply and interactively with the AI tool, leads to better performance, and unstructured streamline, where students used a linear approach with less interaction, results in poorer performance. Consequently, the study concludes that iterative and highly interactive use of AI tools yields better writing results compared to merely using them as supplementary sources.

Escalante et al. (2023) conducted a study to compare the efficacy of AI-generated feedback by using tools like ChatGPT/GPT-4 with that of human tutor feedback on writing. Their research involved 91 English as a New Language (ENL) students in the Asia-Pacific region. The study found no significant difference in learning outcomes between the groups that received AI-generated feedback and those that received human tutor feedback. Interestingly, student preferences were almost evenly split between the two types of feedback. While many students appreciated the clarity and specificity of AI-generated feedback, others valued the personal interaction and engagement offered by human tutors. Consequently, the study suggests that a blended approach which incorporates both AI and human feedback could lead to optimal results in writing instruction.

Similarly, Taiye et al. (2024) examined the impact of Generative AI (GenAI) tools on academic writing and critical thinking skills among 150 undergraduate students and eight teachers in Sweden. The students engaged with GenAI technologies for tasks such as idea generation, summarizing sources, and checking spelling and grammar and some students expressed concerns about potential academic misconduct and the authenticity of their work

while teachers, on the other hand, had mixed reactions. While some were initially resistant to the integration of GenAI due to fears of decreased assignment quality, others were enthusiastic about its potential benefits. Importantly, the study revealed that despite its struggles to create substantive content with evidence, GenAI provided valuable opportunities for discussions about good academic writing practices.

Gayed et al. (2022) conducted a study to evaluate the impact of an AI-based writing assistant on writing performance among ten Japanese EFL learners, and specifically, they focused on the aspects of lexical diversity and fluency. The findings revealed that the use of the AI tool improved lexical diversity, but the improvement was not statistically significant. More notably, the study showed a significant increase in syntactic complexity, and indicated better sentence fluency. This suggests that while AI-based writing assistants can enhance certain aspects of writing, their impact may vary depending on specific metrics and contexts.

As a concluding study, Law (2024) conducted a comprehensive scoping literature review with 41 research studies to assess the current state of research on the application of Generative AI (GenAI) in language teaching. Findings include for instance, five studies specifically examining the use of GenAI for improving writing skills while others looked into automated assessment methods and the ethical implications of using AI in education. Law (2024) noted that while GenAI holds potential for enhancing language learning, it also raises significant ethical concerns such as plagiarism, academic dishonesty, and educational inequality.

In conclusion, the integration of AI in academic writing offers both benefits and challenges. Studies show that AI tools improve writing quality and provide consistent feedback, but their effectiveness varies with student proficiency and interaction. Godwin-Jones (2022) and Alharbi (2023) emphasize the need for teacher involvement to address higher-order skills. On the other hand, Nguyen et al. (2024) found that interactive use of AI like ChatGPT enhances performance while Escalante et al. (2023) advocate for a blended approach of AI and human feedback. As a scoping literature review, Taiye et al. (2024) reveal mixed reactions with concerns about academic misconduct but also benefits for

writing practices. Gayed et al. (2022) note improvements in syntactic complexity and sentence fluency if GenAI is used but limited impact on lexical diversity. Overall, AI tools can enhance writing, but their integration requires careful consideration and human oversight.

2.4. Similar Studies to the Current Study

Given the unique focus of the current study on the comparison the effect of various feedback providers, there is a lack of directly comparable research. However, several related studies offer valuable insights and form the basis for understanding the broader context and potential implications of the current research. Although these studies are not identical in scope or methodology, they share similar characteristics and contribute to the overall discourse on feedback mechanisms in education.

Hamer et al. (2014) conducted a study to compare peer feedback with tutor feedback in an undergraduate software engineering programming class. The study explored differences in the quality, detail, and tone of feedback provided by peers and tutors and the findings revealed that while tutors generally provided longer and more specific comments, peers were equally effective in offering advice. Notably, tutors tended to give more negative feedback whereas peers provided more positive remarks. Interestingly, peer feedback was found to be as beneficial as tutor feedback in certain respects such as advice and personal engagement. Overall, the study highlighted the potential of peer feedback as a valuable complement to tutor feedback, particularly the compensation of tutor feedback in large classes where individual tutor feedback is challenging to provide consistently.

Gielen et al. (2010) investigated whether peer feedback could effectively substitute for teacher feedback and explored various forms of peer feedback to enhance its efficacy. Conducted in a secondary school writing curriculum, the study employed a pre-test post-test control group design to assess long-term learning effects. The findings supported the hypothesis that peer feedback could be as effective as teacher feedback. However, the use of question forms and reply to forms significantly enhanced the benefits of peer feedback.

Almost half of the students found peer feedback helpful, although fewer students felt that providing feedback aided their own learning. The study emphasized the added value of structured peer feedback forms in improving student learning outcomes.

Banihashem et al. (2024) explored the efficacy of AI-generated feedback in comparison to traditional teacher feedback. The study, involving higher education students, highlighted the crucial role of feedback in enhancing learning. Findings indicated that AI-generated feedback could match the quality and effectiveness of teacher feedback in promoting student engagement, motivation, and learning outcomes. The study suggested that AI tools could alleviate the workload on teachers, particularly in large classes, by providing timely and personalized feedback. However, the importance of a blended approach that combines AI-generated feedback with human input was emphasized to maximize the benefits of both methods.

Escalante et al. (2023) examined the impact of AI-generated feedback on writing, focusing on ENL learners. The study involved two longitudinal studies: one comparing learning outcomes between students receiving feedback from ChatGPT and those receiving tutor feedback, and another analysing student perceptions of both feedback types. Results showed no significant difference in learning outcomes between the two groups and the student preferences were nearly evenly split between AI-generated and human feedback. The study concluded that AI-generated feedback could be effectively incorporated into ENL writing instruction without compromising learning outcomes, advocating for a blended feedback approach .

Rollinson (2005) explores the efficacy of peer feedback in ESL writing classrooms and acknowledges both the benefits and challenges associated with its implementation. The study highlights that peer feedback when properly structured and supported through training, can be a valuable complement to teacher feedback. It emphasizes the social, cognitive, and affective benefits of peer interactions, and Rollinson (2005) notes that peer feedback encourages collaborative learning and helps students become more critical readers and writers. Despite initial scepticism from both teachers and students, structured peer feedback

has been shown to lead to significant improvements in writing quality. The study also underscores the importance of creating a supportive environment and providing adequate training to ensure effective peer feedback sessions.

Cui et al. (2022) conducted a comparative study on the effectiveness of peer feedback versus teacher feedback in EFL writing instruction. The research found that while teacher feedback led to more revisions overall, peer feedback fostered greater student autonomy and engagement in the revision process. Teacher feedback was more effective for surface-level corrections such as grammar and spelling whereas peer feedback was more influential in promoting meaning-level changes and enhancing organizational skills. The study concluded that a combination of both feedback types, supplemented by training in providing and utilizing feedback, could yield the best outcomes for EFL learners.

Zhan (2016) investigated the perceptions of written teacher feedback among EFL students and their teacher in a Chinese college setting. The study revealed a discrepancy between the teacher's feedback focus and the students' preferences. While the teacher's feedback covered various aspects including content, organization, vocabulary, grammar, and mechanics, students reported benefiting most from feedback on organization, which was not sufficiently addressed by the teacher. The study suggested that teachers should better align their feedback practices with students' needs and preferences to enhance the effectiveness of writing instruction. Improved communication between teachers and students about feedback expectations was recommended in the study.

Qi et al. (2024) examined the impact of feedback from both teachers and ChatGPT on students' learning outcomes in higher education. The study, based on a sample of 300 students in computing-related subjects, found that feedback from both sources significantly improved self-efficacy and affective engagement, and this in turn positively influenced learning outcomes. Teacher feedback was more impactful on students' self-efficacy while ChatGPT's feedback was particularly effective in enhancing affective engagement due to its instant and personalized nature. The study recommended the hybrid use of both types of feedback to optimize student learning.

Montgomery and Baker (2007) examined the alignment between teacher self-assessments, student perceptions, and actual teacher feedback in L2 writing. The study revealed a gap between teachers' self-perceptions of their feedback practices and students' preferences. While teachers believed they provided balanced feedback on both local and global issues, students felt that the feedback was predominantly focused on local issues such as grammar and mechanics. The findings emphasized the need for teachers to consider student feedback preferences and to engage in reflective practices to better align their feedback with student needs.

These studies, although they are not identical to the current research, they provide valuable insights into the dynamics of feedback in educational settings, highlighting the significance of combining different feedback sources and aligning feedback practices with student preferences to enhance learning outcomes.

2.5. The Summary of the Chapter

Chapter Two of the study reviews literature on teaching writing, feedback mechanisms, and GenAI in academia. It explored different approaches to teaching writing, including product, process, and genre approaches. Each of them emphasizes structural mastery, developmental processes, and contextual understanding. Furthermore, the feedback mechanisms were categorized into teacher and peer feedback, delivered face-to-face or online. Effective feedback was highlighted as clear, specific, timely, and balanced, and it should foster linguistic accuracy and overall competence. The AMMW model by Razi was introduced as an innovative approach to enhance writing through anonymous and diverse feedback in process writing approach. The model incorporates scaffolding, the ZPD, and metacognition, and it also demonstrates its effectiveness in improving writing skills and promotes learner autonomy.

The chapter also covered the evolution and application of GenAI in writing and traces significant milestones in AI history. It underscored the importance of prompt engineering in optimizing LLMs to maximize their potential in educational settings. Additionally, several

related studies provided valuable insights into the broader context of feedback mechanisms in education. Although these studies were not identical in scope or methodology of the current research, they highlighted the potential benefits and challenges of different feedback sources, including peer, teacher, and AI-generated feedback.

Furthermore, the GenAI section of Chapter Two provided a comprehensive overview of the advancements, applications, and ethical implications of GenAI in writing, and emphasized its transformative potential and the necessity of responsible integration in educational settings. Overall, Chapter Two provided insights into teaching writing, effective feedback, and integration of GenAI, and presented important information for educators and researchers.

CHAPTER THREE

METHODOLOGY

This chapter explains the research methodology employed in this study, and thoroughly presents the details about the design, setting, participants, data collection tools, and data analysis procedures. The chapter begins with a detailed description of the research design, which is followed by the specific research questions the study is aimed to shed light on. The setting and participants are then described, covering the structure of the course and the selection procedures for participants. The data collection tools including the course assignments, the Fundamentals of Academic Writing Rubric, GenAI prompts, and interview questions are presented with details. Following sections cover the data collection procedures including the data collection of the assignments and feedback of the assignments and participant interviews. Finally, the data analysis procedures for each research question are outlined, and the chapter is concluded in a brief summary of the chapter.

3.1 . Research Design

In this study, the primary aim is to measure the effects of GenAI feedback compared to other feedback providers including teachers and peers. Additionally, it aimed to determine if there were significant differences among them. Finally, the strengths and weaknesses of these feedback providers were investigated. Accordingly, this research adopted the mixed method approach with instrumental collective case study design. As the research was conducted on a specific sample to evaluate the effects of feedback on and participants' perceptions of these feedback types, it was regarded as a case study, and as the research also aimed to illuminate an issue benefitting from the cases, it was also regarded as an instrumental study. In conclusion, this study utilized a collective instrumental case study design to effectively compare and analyse the impacts, strengths, and weaknesses of GenAI, teacher, and peer feedback.

Regarding the approach of the research, the mixed method was adopted since the researcher aimed to address the questions within a pragmatic framework, and consequently

gathered both quantitative and qualitative data in a sequential manner to obtain the necessary answers (Creswell, 2002). In the mixed method approach, Creswell (2002) identifies two phases of mixed method approach including the collection of different types of data and the collection of quantitative data to obtain detailed information from the participants. In that way, the blend of both methods including qualitative and quantitative serves the aim of creating a constructive interaction to complement each other with the aim of developing a theory or illuminating a complex issue (Dörnyei, 2007). In conclusion, the mixed method approach enabled the researcher to comprehensively address the research questions by utilizing the strengths of both qualitative and quantitative data to develop a deeper understanding and provide a more thorough analysis.

Regarding the design, collective instrumental case study design was adopted. Case studies, according to Stake (1995), are divided into three categories including the intrinsic case study, instrumental case study, and multiple or collective case study. Stake (1995) defines an intrinsic case study as research conducted by someone motivated to learn about the case itself, whereas an instrumental case study is defined as research where the primary case being studied helps illuminate a secondary case. In the light of these explanations, Dörnyei (2007) regards collective case study as an instrumental case study since collective case studies also cover other cases to understand one another. In collective case studies, rather than using a case to learn about the issue, several cases could be co-ordinately used to learn about the issue (Stake, 1995). Stake (2006) refers collective case as a multicase study and stated that single cases are given importance because each case belong to a collection of cases that are the part of the whole, and these cases are interconnected. The aim is to understand the entirety of these connected cases and in order to achieve this, each case is studied individually rather than as a whole, and each is evaluated separately to understand the complete picture (Stake, 2006). In conclusion, the collective instrumental case study design was adopted to understand interconnected cases. By studying each case individually, the approach aimed to provide deeper insights into the complete picture, highlighting the value of each case as part of the whole.

3.2. Research Questions

The research questions of the current study are threefold including:

1. Is there a significant difference between:

1.1. the gain scores of feedback provided by teachers, peers, and ChatGPT on students' essay scores?

1.2. the pre-scores given by the teacher and the scores given by peers and GenAI?

2. How does the feedback provided by the teacher, peers, and artificial intelligence affect the improvement of students' assignment grades, as measured by pre- and post-feedback scores?

3. How do students perceive the strengths and weaknesses of feedback provided by teacher, peers, and GenAI?

3.3. Setting and Participants

This study was conducted with 16 students out of 132 students who successfully completed Writing Skills I course during the fall semester of the first year at the Department of English Language Teaching, Faculty of Education, at a state university in Türkiye. This department and course were chosen in line with the aims and research questions of the study because writing skills, which were the one of the cornerstones of this study, formed the foundation of this course. Additionally, the course instructor's existing interest in integrating artificial intelligence was also utilized in this study. In this way, an authentic and organic setting existed for conducting the study. In the following sections, details about the course and information regarding the selection and elimination of participants were presented sequentially.

3.3.1. Structure of the Course

The following information about the course was obtained from the instructor and detailed in the Appendix A. Officially titled as ‘Writing Skills I’, the course spanned 14 weeks with 2 hours of instruction per week. The language of instruction was English, and the teaching methods included lecturing, extensive reading, and feedback exchanges between students, GenAI and the teacher. The course instructor developed this structure to provide a comprehensive learning experience for students. The course aimed to develop academic writing skills through critical analysis of texts and readings, enhancing students’ vocabulary knowledge. It emphasized writing academic texts, exchanging feedback, and learning the APA (American Psychological Association) citation rules. One of the course outcomes was for students to present and defend their viewpoints clearly and persuasively in an academic setting.

Assessment involved three essay assignments, each was 700-word-long essays on average and the essay writing process included the drafting, receiving feedback, revising, and resubmitting stages. Feedback was provided by the teacher on the first assignment, peers on the second assignment, and GenAI on the third assignment. The feedback and the assessment rubric were conducted by using the Fundamentals of Academic Writing Rubric (FAWR) (Razi, 2023), Salim Razi, and FAWR was explained in detail in the data collection tools section in this chapter.

To keep track of the feedback received by students, it was required for students to upload revised versions of their assignments to Turnitin after each feedback session. This process was regarded as an essential step for assessing how effective the feedback was and how students were applying it. For the first assignment, the teacher provided feedback via Turnitin using FAWR (Razi, 2023) as mentioned before. After receiving teacher feedback, students were required to resubmit their assignments showing the changes made using the track-change feature of Microsoft Word, along with a version without track-changes.

For the second assignment, students also used FAWR on Turnitin Feedback Studio to exchange feedback with peers. The peer feedback process involved providing feedback to three peers and receiving feedback from three peers, with assignments made randomly and anonymously by the system. Therefore, students did not know who provided or received their feedback. After receiving peer feedback, students were required to resubmit their assignments showing the changes made using the track-change feature of Microsoft Word, along with a version without track-changes, as well.

For the third assignment, students used GenAI, specifically ChatGPT 3.5, to manage feedback using prompts developed according to FAWR (Razi, 2023). In the case of ChatGPT feedback sessions, students were required to export their chat logs, which were then uploaded to cloud storage applications and shared with the teacher and peers. This allowed students to review each other's ChatGPT feedback sessions ensuring that students not only received feedback but also engaged actively in the GenAI feedback of each other. The development of these prompts was detailed in the data collection tools section. After receiving feedback, students were required to resubmit their assignments showing the changes made using the track-change feature of Microsoft Word, along with a version without track-changes.

Assessment of these assignments involved several key components including the quality and accuracy of the peer reviews provided by students, with each student submitting a total of three peer reviews. Additionally, the appropriate and ethical use of GenAI tools, were assessed by examining how accurately and ethically students received, applied, and managed the feedback from GenAI in their assignments. Feedback sessions were counted as a mid-term score forming 40% of the term grade, while the writing tasks counted as the final examination forming 60% of the term grade. The Figure 1 demonstrates the grading breakdown for the Writing Skills I Course. It shows how the total score out of 100 is divided between mid-term and final components, with further breakdowns of percentages allocated to various feedback sessions and assignments. The course attendance policy required students to attend 70% of the total course hours. The academic integrity policy was strictly enforced, emphasizing the seriousness of plagiarism and unethical use of AI. The instructor stated that students involved in plagiarism, collusion, contract cheating, or unethical use of

AI would receive zero points. The instructor meticulously outlined procedures for registration, assignment submission, and feedback processes within Turnitin, offering detailed guidelines encompassing submission procedures, GenAI prompts, and examples of effective and ineffective feedback. Furthermore, the instructor provided comprehensive information on course prerequisites, instructional resources, assignment specifications including document formatting, primary and supplementary reading materials, and weekly discussion themes, including assigned readings and tasks for each session.

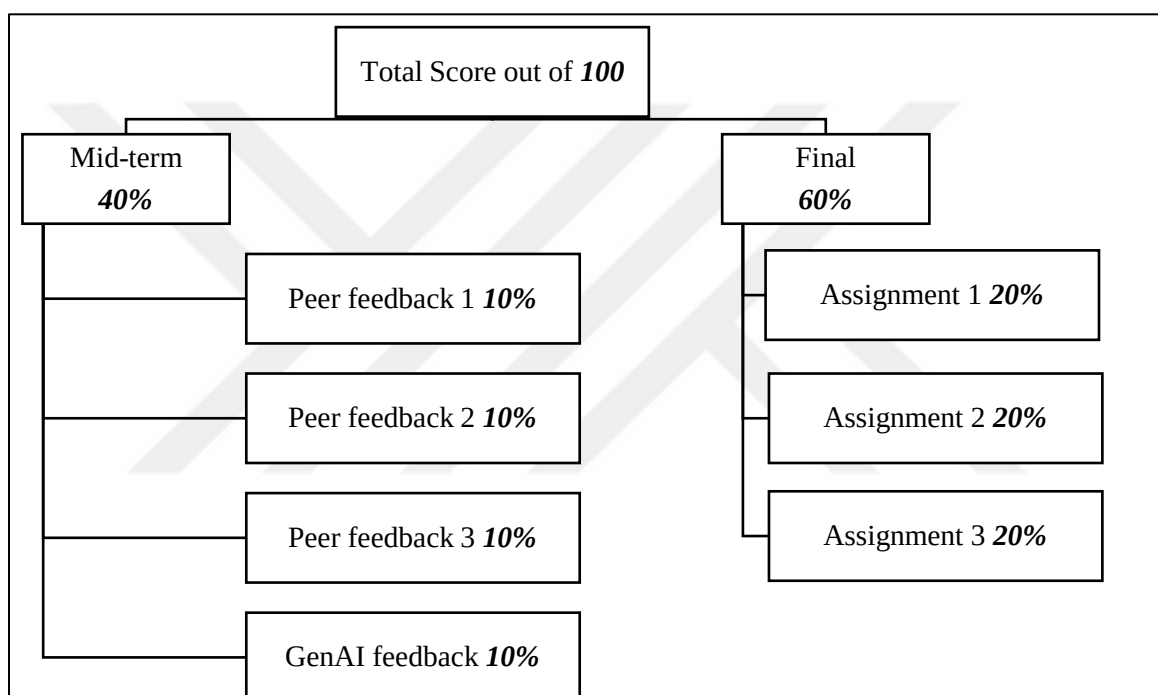


Figure 1. Grading Structure Diagram for the Writing Skills I Course

3.3.2. Selection of Participants

The selection of participants was guided by Stake's (2005) recommendation that multicase studies are optimally studied with a limited number of cases, typically between four and ten, as fewer than four cases may lack sufficient interaction for analysis, while more than ten cases may present challenges in effectively examining their interactions. Stake (2005) also noted that many multicase studies include more than 15 cases for justified reasons. In this study, 16 participants were selected for analysis. The selection process involved several methodical phases.

The first phase started with the course instructor completing fall semester grading for 132 enrolled students. The grading of seven components included Assignment 1's paper, three peer reviews for Assignment 2, Assignment 2's paper, utilization of GenAI feedback for Assignment 3, and Assignment 3's paper. Subsequently, the first phase continued with the elimination of the students who did not fulfil all assigned tasks and attendance requirements, resulting in a cohort of 103 students who successfully met the criteria for participation in the Writing Skills I course. In the following part of the first phase, the mean total score for these 103 students was calculated ($M=78.06$, $N=103$), with scores ranging from a minimum of 56.50 to a maximum of 95. Then, students were categorized into those scoring below the mean ($n=45$) and above the mean ($n=58$).

Table 1

Selected Participants' Scores with Mean, Minimum, and Maximum Comparisons

Participant No	Total Score	Condition
P1	75.1	$\leq$ Mean (78.06)
P2	88.9	$\geq$ Mean (78.06)
P3	75.6	$\leq$ Mean (78.06)
P4	87.3	$\geq$ Mean (78.06)
P5	89	$\geq$ Mean (78.06)
P6	67.5	$\leq$ Mean (78.06)
P7	75.7	$\leq$ Mean (78.06)
P8	94.5	$\geq$ Mean (78.06)
P9	95	= Maximum (95)
P10	89	$\geq$ Mean (78.06)
P11	89.3	$\geq$ Mean (78.06)
P12	90.1	$\geq$ Mean (78.06)
P13	71.7	$\leq$ Mean (78.06)
P14	57.5	= Minimum (57.5)
P15	89.6	$\geq$ Mean (78.06)
P16	72.1	$\leq$ Mean (78.06)

In the second phase of participant selection, 16 students were chosen based on the literature's recommendation that multicase studies should ideally include up to 15 participants, which constituted approximately 15% of the research population of 103 students. These participants were subsequently categorized into two groups: those scoring below and those scoring above the mean value of the teacher's scores ($M=78.06$). The selected participants' scores, ranging from a minimum of 57.5 to a maximum of 95, were

detailed in Table 1 above. Out of the 16 participants, 9 scored above the mean and 7 scored below it, reflecting the same ratio observed in the larger group of 103 students, where 45 scored below the mean and 58 above it. This sampling strategy aimed to accurately represent the broader student population, with 15% of 45 and 15% of 58 participants selected, totalling 16 participants. Figure 2 visually outlines the participant selection process, starting from 132 initial participants, filtering down to 103 successful students meeting course criteria, dividing them into groups based on mean scores, and ultimately selecting 7 and 9 participants respectively from each group.

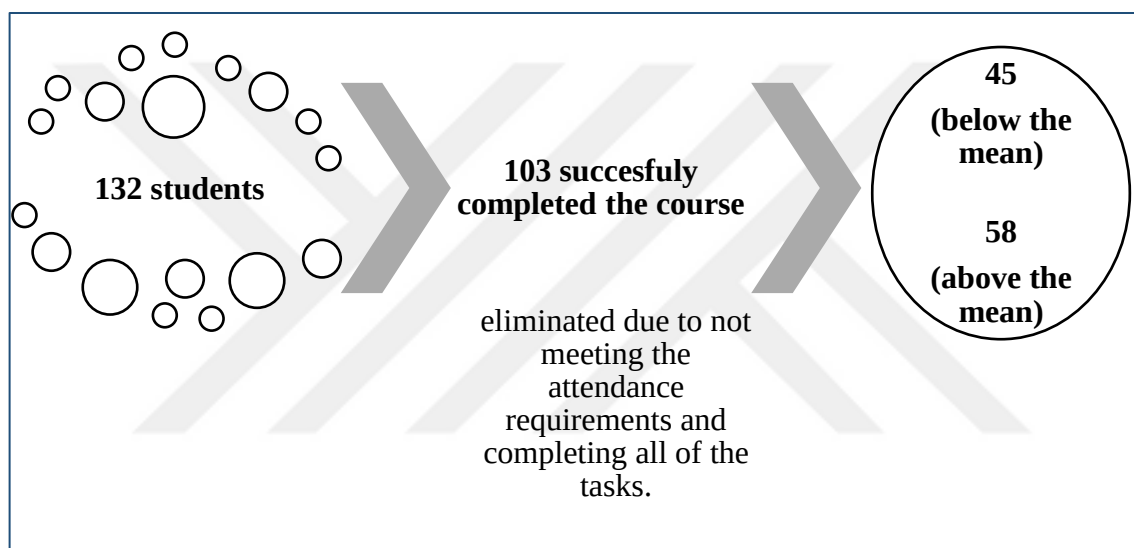


Figure 2. Participant Selection Process.

3.4. Data Collection Tools and Data Collection Sources

Data collection for this study employed a comprehensive array of tools and sources encompassing both quantitative and qualitative methodologies consisting of two data collection tools and two data collection sources. Initially, data were gathered through the course assignments forming the first data collection sources, which consisted of three essay-writing tasks and feedback sessions. Subsequently, the assessment of these assignments was conducted using the Fundamentals of Academic Writing Rubric (FAWR), serving as the first data collection tool. The second data collection source involved GenAI prompts developed in accordance with FAWR criteria by the researcher. Finally, qualitative insights were obtained through interview questions formulated by the researcher. This section provides

detailed explanations of each of these data collection tools and sources. Starting from the introduction of FAWR, continuing with the assignments of the course and the development of the prompts, Data Collection Tools and Sources section was finalised with the narration of interview questions’ development. Table 2 summarizes the key components of the research methodology outlining the research questions addressed, the sources of data used, the methodology employed, the data collection tools utilized, and finally, the data analysis procedures.

Table 2

Key Components of the Research Methodology

Research Question	Data Collection Sources	Methodology	Data collection tools	Data analysis
1.1.	Gain scores of Assignment 1, 2, and 3	Quantitative	FAWR	Non-parametric Wilcoxon Signed Ranks Test
1.2.	Pre-scores of the Assignment 2 and 3 from the teacher and the feedback scores from peers and GenAI	Quantitative	FAWR	Non-parametric Wilcoxon Signed Ranks Test
2	Pre- and Post-scores of Assignment 1,2 and 3	Qualitative and Quantitative	FAWR	Narration and Descriptive Statistics
3	Participants	Qualitative	Interview questions	Narration and Qualitative Content Analysis

3.4.1. Fundamentals of Academic Writing Rubric as a Data Collection Tool

The Fundamentals of Academic Writing Rubric (FAWR) by Razı (2023) serves as a more straightforward alternative to the Transparent Academic Writing Rubric (TAWR) (Razı, 2015). TAWR comprises 50 items categorized under introduction, citations, academic writing, idea presentation, and mechanics, with scoring options of zero for poor, one for acceptable, and two for excellent. Razı (2015) established the construct validity of TAWR,

including discriminant and convergent validity, using a multi-trait multi-method matrix. Similarly, FAWR encompasses five categories including introduction, citations, academic writing, idea presentation, and mechanics but reduces the total number of items to 20, each with five scoring options ranging from zero to five. For this study, however, the FAWR scoring system was adapted to include scores of zero, 2.50, and five, with detailed explanations provided by the course instructor to the researcher. The adapted FAWR format is detailed in Appendix B. In conclusion, FAWR maintains the essential categories and rigor of TAWR while offering a simplified and modified scoring system to enhance clarity and consistency in academic writing assessments.

3.4.2. Assignments of the Course as Data Collection Sources

During the semester, students were required to submit three essays, each approximately 700 words in length, including references, on distinct topics. The first essay focused on the use of L1 in EFL education for young learners, the second on strategies to promote extensive and intensive reading among EFL young learners, and the third on enhancing listening skills in the same group. These assignments served not only as academic tasks but also as data collection sources for the study. They were evaluated using the Fundamentals of Academic Writing Rubric (FAWR), which was detailed in the previous section. Evaluation of these assignments provided data for the researcher to determine pre-scores, post-scores, and gain scores for addressing Research Questions 1 and 2. Each assignment also included a feedback session including feedback for Assignment 1 from the teacher, Assignment 2 from the peers, and Assignment 3 from GenAI. These feedback sessions were structured upon FAWR, ensuring consistency in assessment and feedback processes throughout the study.

3.4.3. Development of the Prompts as Data Collection Sources

The development of GenAI prompts required researcher to conduct an extensive study of prompt engineering, a concept discussed in the literature review. This comprehensive study involved reviewing pertinent research studies, enrolling in MOOCs,

viewing instructional videos, actively participating in GenAI forums, engaging in practical experimentation with GenAI, and consulting with experts and colleagues in the field. Consequently, the researcher generated prompts through a multi-stage process, including conceptualization, prompt formation, question refinement, pilot testing, and evaluation. The developed prompts are presented in the Appendix C. This iterative process could extend beyond five stages if the prompt required further refinement after initial evaluations, based on the FAWR items.

In the initial stage, conceptualization employed the context manager prompt pattern under the category of context control, as delineated by White et al. (2023). The context manager pattern aims to delineate the issue's boundaries for AI conversations, focusing on specific aspects without irrelevant information. This was achieved by querying GenAI on its understanding of the FAWR items to include or exclude relevant contexts. An example of this stage was illustrated in Figure 3 below.

Development of the prompt – Stage 1 – Getting ready for the context manager pattern

Human: Provide guidelines for evaluating the effectiveness of a topic introduction in an essay.

GenAI: To evaluate the effectiveness of a topic introduction in an essay, consider the following guidelines:

1. **Clarity of Purpose**: The introduction should clearly state the purpose or thesis of the essay. Ensure that readers understand what the essay intends to discuss or argue.
2. **Engagement**: A compelling introduction grabs the reader's attention. You can achieve this through a thought-provoking question, an interesting fact, a relevant anecdote, or a powerful quote.
3. **Context**: Provide context for your topic. Explain why it's important, relevant, or timely. Help the reader understand why they should care about the subject matter.
4. **Thesis Statement**: The introduction should contain a concise and clear thesis statement that outlines the main argument or point of the essay.
5. **Organization**: Ensure that the introduction sets the stage for the rest of the essay. It should give readers an idea of how the essay will be structured.
6. **Transition**: Use transitional words or phrases to smoothly connect the introduction to the body paragraphs. This helps the essay flow logically.
7. **Length**: Keep the introduction concise. It should be long enough to convey necessary information but not so lengthy that it becomes overwhelming.
8. **Avoiding Plagiarism**: Make sure the introduction is original and not copied from other sources. Properly cite any external information or quotes.
9. **Relevance**: Ensure that every sentence in the introduction is directly related to the topic and thesis. Avoid irrelevant information or tangents.
10. **Revising and Proofreading**: After writing the introduction, revise it for clarity, coherence, and grammar. Eliminate any errors that may distract the reader.
11. **Audience Awareness**: Consider your target audience and tailor the introduction to their level of knowledge and interest in the topic.
12. **Feedback**: Seek feedback from peers, teachers, or writing tutors. Others' perspectives can help you improve your introduction further. By following these guidelines, you can create a strong and effective introduction for your essay.

Figure 3. Development of the prompt – Stage 1 – Getting ready for the context manager pattern.

During the second stage, the guidelines provided by GenAI were meticulously inspected to manage its understanding of the items of the rubric, such as topic introduction. The researcher used contextual statements like ‘within scope,’ ‘considering,’ ‘ignoring’ and ‘in terms of’ to refine the prompt’s boundaries, as outlined by White et al. (2023). An example of this stage was shown in Figure 4.

Formation of the prompt – Stage 2 – Formation of the prompt with the context manager pattern

Asses the introduction of this essay in terms of its ability to engage the reader, provide a clear overview of the topic, and establish the context for the discussion. Provide specific feedback with numbered points, substantiated by evidence and examples from the text. Additionally, rate the topic introduction with one score (integer) on a scale of 0 to 2 based on the following criteria:

0 for no introduction of the topic or problem

1 for an inadequate introduction of the topic

2 for an effective introduction

Figure 4. Formation of the prompt – Stage 2 – Formation of the prompt with the context manager pattern.

In the third stage, the question refinement pattern was applied to the prompt to make it better. This pattern, under the prompt improvement category, guided GenAI to create a better version of the prompt, enabling quicker, more accurate responses. Contextual statements such as ‘Within scope X, suggest a better version of the question to use instead’ were utilized (White et al., 2023, p. 8). An example of this stage was presented in Figure 5.

Question Refinement Pattern – Stage 3

Human: Suggest a better version of the prompt below without removing any information.

Asses the introduction of this essay in terms of its ability to engage the reader, provide a clear overview of the topic, and establish the context for the discussion. Provide specific feedback with numbered points, substantiated by evidence and examples from the text. Additionally, rate the topic introduction with one score (integer) on a scale of 0 to 2 based on the following criteria:

0 for no introduction of the topic or problem

1 for an inadequate introduction of the topic

2 for an effective introduction

GenAI: Evaluate the essay’s introduction for its capacity to captivate the reader, present a concise overview of the subject matter, and set the groundwork for ensuing discussion. Offer precise feedback utilizing numbered observations, supported by evidence and instances from the text. Also, assign a single score (whole number) on a scale from 0 to 2 based on the subsequent criteria:

0 for the absence of an introduction to the topic or issue

1 for an insufficient introduction,

2 for a proficient introduction.

Figure 5. Question refinement pattern – Stage 3.

The fourth stage involved piloting the prompt on five different student essays. Before piloting, ChatGPT 3.5 was customized with specific instructions, including user information such as profession and hobbies, and behavioural preferences for ChatGPT. These custom instructions were detailed in Figure 6. After setting these instructions, five essays, provided by the course instructor, were tested. The essays represented a range of student performance, with approximately scores of 20, 40, 60, 80, and 100 out of 100.

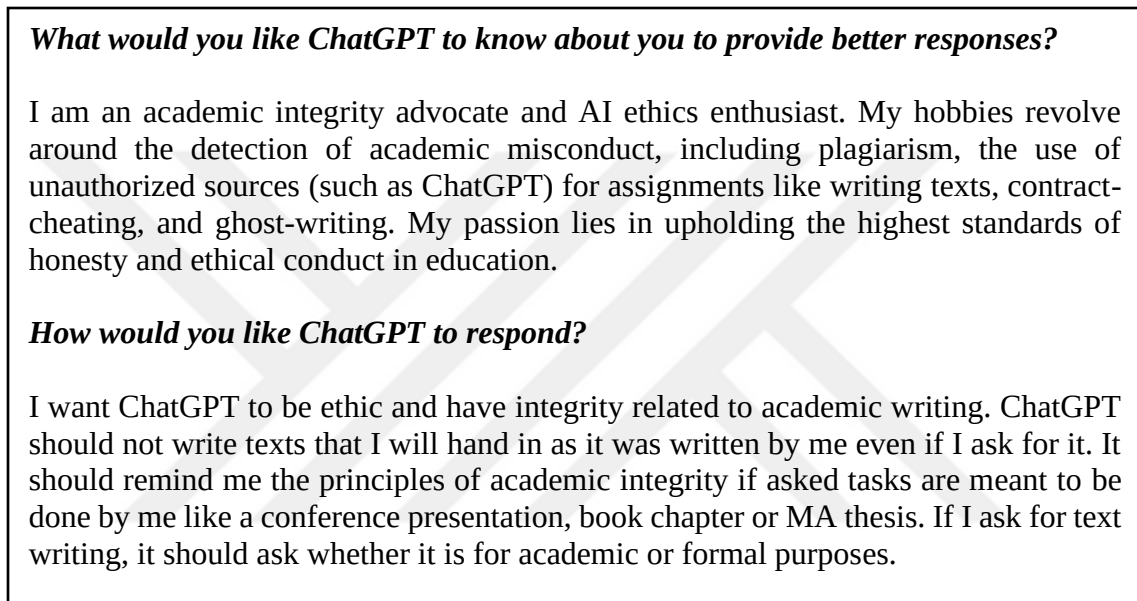


Figure 6. Custom Instructions for ChatGPT.

In the fifth stage, the success of the pilot essays was evaluated. Each essay was assessed based on several criteria:

- Is there an evaluation for captivating the reader?
- Is there an evaluation for the overview of the subject matter?
- Is there an evaluation for setting the groundwork for ensuing discussion?
- Are these evaluations presented in numbered points, supported by evidence and instances from the essay?
- Is there a single score assigned to the essay's introduction with explanations?

If these criteria were satisfactorily met, the prompt was regarded ready for use. If not, the prompt was refined, starting again from the second stage. For instance, in the fourth stage which was piloting the prompts, ChatGPT initially assigned a score to the entire essay rather than just the introduction. After encountering this issue in over half of the pilot tests, the prompt was refined with additional specification words and retested with five essays. This procedure was repeated for all 20 FAWR items to ensure comprehensive prompt development.

3.4.4. Development of Interview Questions as a Data Collection Tool

To evaluate the strengths and weaknesses of different feedback providers, six interview questions were developed. These questions aimed to gather insights from participants regarding teacher feedback, peer feedback, and ChatGPT feedback. To ensure participants could express their thoughts accurately and genuinely, the questions were formulated in their native language, Turkish. The researcher believed that participants would articulate their ideas more effectively in their native language. Expert opinion from the course instructor was received from the course instructor to ensure the reliability of the interview questions. In conclusion, the interview questions were meticulously designed and validated to gather comprehensive insights on the strengths and weaknesses of teacher, peer, and ChatGPT feedback from participants in their native language. The interview questions were as follows:

- 1. According to your experience, what were the strengths of the teacher feedback?*
- 2. According to your experience, what were the weaknesses of the teacher feedback?*
- 3. According to your experience, what were the strengths of the peer feedback?*
- 4. According to your experience, what were the weaknesses of the peer feedback?*
- 5. According to your experience, what were the strengths of ChatGPT feedback?*
- 6. According to your experience, what were the weaknesses of ChatGPT feedback?*

3.5. Data Collection Procedures

Equally important as the data collection tools and sources discussed in the previous section are the data collection procedures, which hold significant importance for the current research. This section provides an explanation of the data collection procedures for the assignments, detailing the roles of the feedback providers, as well as the interviews conducted with the participants.

3.5.1. Data Collection of the Assignment 1 and the Teacher Feedback

The data collection process for Assignment 1 and subsequent teacher feedback involved two datasets: pre-scoring and post-scoring of the assignment. For the submission of all the assignments and feedback sessions, all students were required to enrol in a Turnitin Classroom created by the course instructor. Students submitted their assignments to Turnitin, which enabled the instructor to provide feedback and assign scores for the Assignment 1. For this study, the researcher selected participants who consented to join and created an Excel document for each participant, including the rubric. The researcher initially scored each item as a pre-score based on the papers downloaded from Turnitin. After receiving the instructor's feedback, students were required to upload revised versions of their assignments, both with and without track changes. The researcher then downloaded these revised assignments and created a separate Excel document for each participant to score the assignments as a post-score. In conclusion, both pre-scores and post-scores were systematically collected for each participant on Assignment 1, providing a comprehensive dataset for analysis.

3.5.2. Data Collection of the Assignment 2 and Peer Feedback

For Assignment 2 and peer feedback, a total of five datasets were collected. The students, who had previously enrolled in the Turnitin Classroom created by the course instructor, used the same classroom to submit their second assignments. For this assignment, similar to the first assignment, the researcher created an Excel document for each participant,

including the rubric, and initially provided pre-scores. As the peers provided feedback using the rubric, an Excel document was created for each of the three peers, recording the scores assigned to each item. Thus, three additional datasets were generated. After receiving peer feedback, students uploaded revised versions of their assignments to Turnitin, both with and without track changes. The researcher then scored these revised assignments as post-scores, using separate Excel documents with the rubric. In conclusion, a total of five datasets were collected: pre-scores, post-scores, and three datasets for peer feedback. This comprehensive data collection facilitated a detailed analysis of the assignment and feedback process.

3.5.3. Data Collection of the Assignment 3 and GenAI Feedback

For Assignment 3 and GenAI feedback, three datasets were collected. The students submitted their third assignments to the Turnitin Classroom. Subsequently, the researcher conducted a two-week, four-hour prompt training session. During this training, the researcher discussed her thesis, provided brief information about GenAI and large language models, summarized the feedback sessions conducted so far, and explained how she developed prompts. The researcher also demonstrated custom instructions and their implementation, applying the developed prompts live in the classroom with the students. The responses from ChatGPT were evaluated together with the students, focusing on how to assess feedback, distinguish correct from incorrect or irrelevant responses, and use the feedback ethically without committing academic misconduct. The researcher shared a document outlining the steps for obtaining feedback from ChatGPT, starting from creating a ChatGPT account, adding custom instructions, naming the chats, exporting chat transcripts, and where and how to upload these exports, including the required prompt document. Although all these details were covered during the face-to-face training, the document was also shared as a reference. After students shared their ChatGPT exports in the specified format, the researcher downloaded the selected students' exports and entered their ChatGPT scores into separate Excel documents. Then, students uploaded the final versions of their Assignment 3, showing revisions with and without track changes, to Turnitin. The researcher downloaded these final assignments and assigned post-scores in Excel documents for the selected participants. In summary, three datasets were collected for this process: pre-scores,

post-scores, and ChatGPT scores, facilitating a comprehensive analysis of Assignment 3 and the GenAI feedback process.

3.5.4. Interviews with the Participants

In the interviews conducted with participants, their native language, Turkish, was used. Six questions were asked during the interviews, which lasted an average of 10 minutes. These interviews were conducted via Zoom and recorded. Prior to the interviews, participants were informed that their information would be anonymized and would not be used for any purpose other than this thesis. They were asked if they wished to proceed with the interview and were informed that the study had received ethical approval for data collection. After participants consented to participate, the questions were asked, and all interviews were transcribed. The transcriptions were then translated into English and sent back to the participants for verification. Once participants approved the English translations of their responses, the transcriptions were analysed, as discussed in subsequent sections. In total, 16 transcription files were created from the interviews, each averaging 15 pages.

3.6. Data Analysis

Regarding the data analysis of the study, qualitative and quantitative analysis were conducted on the data collected from the sources. Analysis procedures include the employment of statistical tests, narration, and qualitative content analysis. In this section, data analysis of each research question were explained in detail consisting of three sections.

3.6.1. Analysis Procedures for Research Question 1

Under Research Question 1, the study investigated two main areas. The first area explored whether there were the significant differences in gain scores of feedback provided by teachers, peers, and ChatGPT on students' essay scores. The second area examined whether there was a significant difference between the pre-scores given by the teacher and the scores given by peers and GenAI.

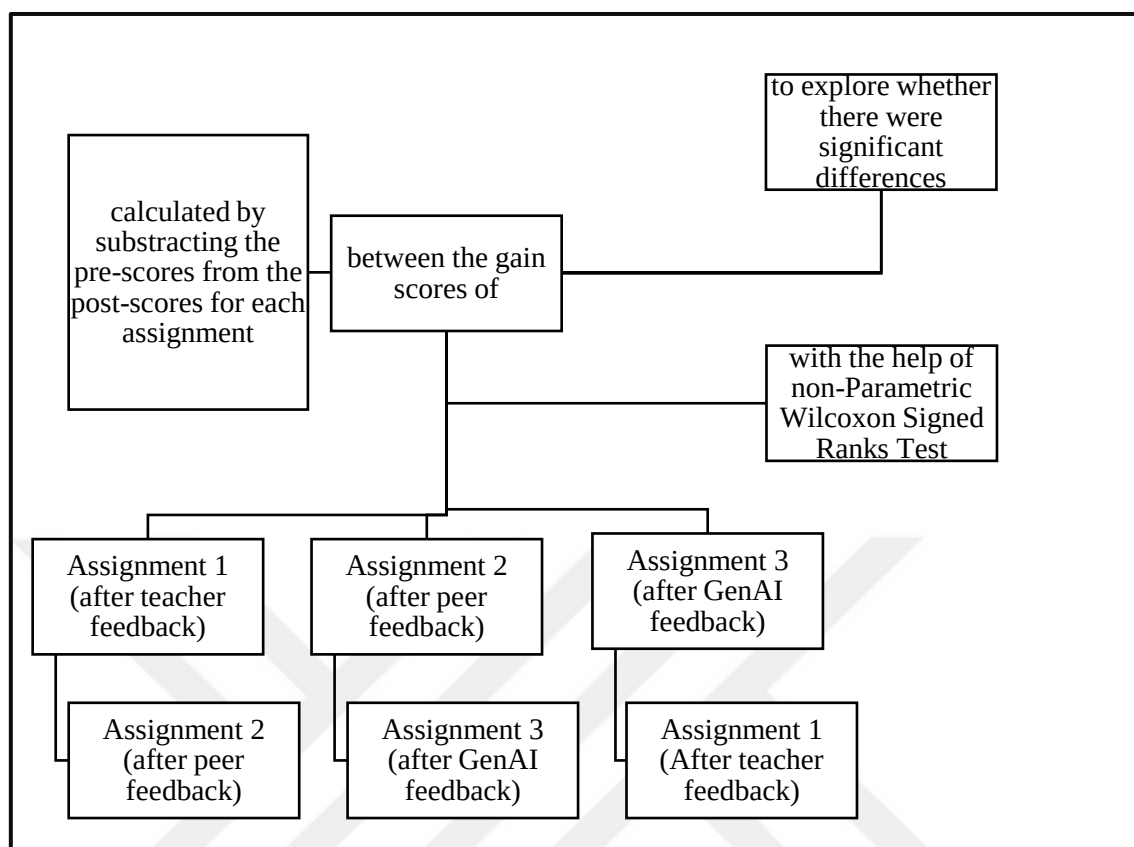


Figure 7. Data Analysis Procedures for Research Question 1.1.

For the first area, the gain scores of the selected participants were calculated. This involved subtracting the pre-scores from the post-scores for each of the three assignments separately to determine the gain scores. These gain scores were then compared using the non-parametric Wilcoxon signed ranks test in SPSS, the direction two ways of comparisons include the comparison of the gain score of Assignment 2 and Assignment 1, gain score of Assignment 3 and Assignment 1, gain score of Assignment 3 vs. Assignment 2. The aim of the researcher with the research question 1.1. was to determine the effectiveness of feedback from different sources. By comparing the gain scores, which were the improvements in students' essay scores from pre-score to post-score, the study was aimed to assess whether feedback from teachers, peers, or ChatGPT leads to significantly different levels of student improvement. This could provide insights into which type of feedback was the most beneficial for enhancing students' writing skills. The diagram for the analysis procedures of the Research Question 1.1. was presented in the Figure 7 above.

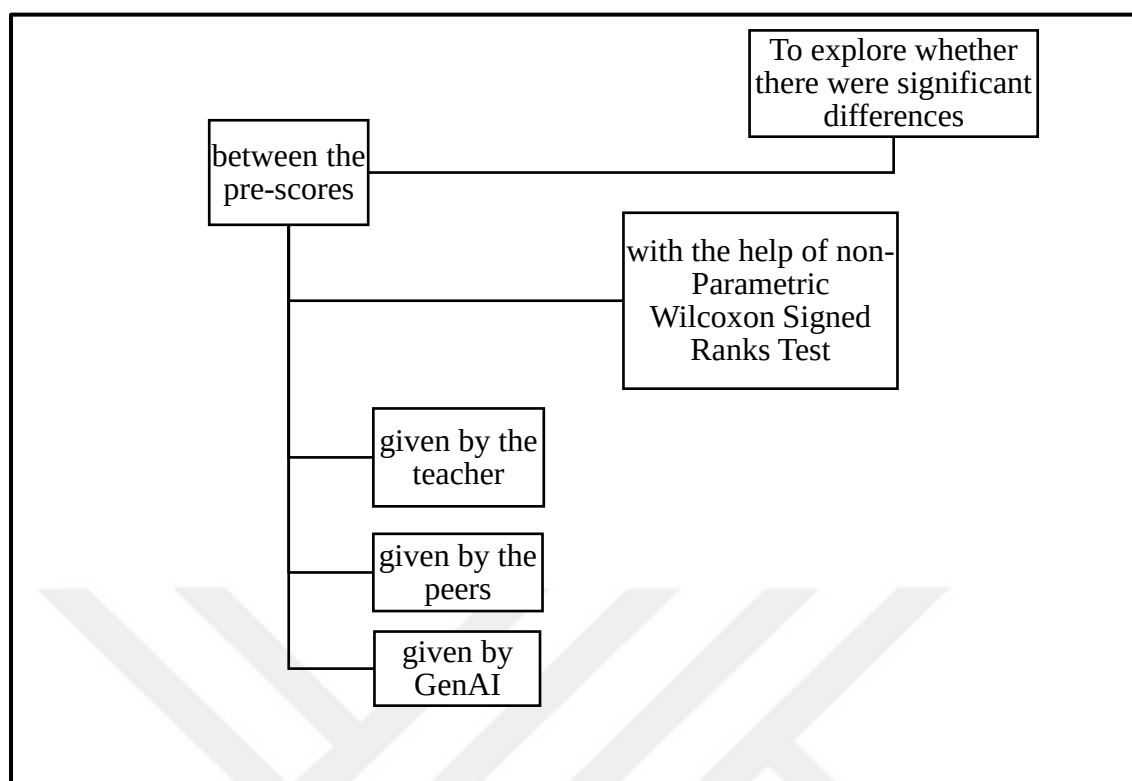


Figure 8. Data Analysis Procedures for Research Question 1.2.

For the second area, the pre-scores of participants from Assignment 2 and Assignment 3 were compared with the scores given by peers and GenAI. This comparison was also conducted using the non-parametric Wilcoxon signed ranks test. The aim of researcher with the research question 1.2. was to determine whether there was a statistically significant difference in the initial evaluation of student assignments by different evaluators. Specifically, the researcher aimed to compare the consistency and accuracy of the pre-scores given by the teacher with those provided by peers and GenAI, assessing if these different sources of feedback produce similar or varying initial scores. For the peer scores, the average was calculated and used for comparison. Although peer feedback was required to be obtained from three peers, in some cases, two peers only provided feedback. In such instances, the average of the two peer scores was used for comparison. This could help in understanding the reliability and validity of peer and AI evaluations compared to traditional teacher assessments. The diagram for the analysis procedures of the Research Question 1.2. was presented in the Figure 8 above.

3.6.2. Analysis Procedures for Research Question 2

Regarding the analysis procedures for the Research Question 2, the study was aimed to investigate the effects of feedback from the teacher, peers and GenAI measured by pre- and post-feedback scores. Initially, individual tables were compiled for each participant that detailed the scores for each assignment and the feedback provided. These tables captured the pre-feedback scores, the specific feedback for each rubric element, and the post-feedback scores, allowing the researcher to identify which elements were revised and the corresponding score improvements. Qualitative narration techniques were used to describe the detailed feedback process for each participant involving the examination of the individual tables to narrate which specific rubric elements were addressed in the feedback, how the feedback was implemented by the participants, and the resultant score improvements. Then, based on the effect of feedback from teacher, peers and GenAI, participants' scores, treated as cases, were explained in detail in the following sections. This qualitative analysis provided a comprehensive understanding of the feedback's impact on specific aspects of the participants' work.

Secondly, the researcher analysed the average performance improvements of 16 participants after receiving feedback on three different assignments. Each participant received feedback from a teacher, peers, and GenAI. For the first assignment, participants received teacher feedback, for the second assignment, peer feedback, and for the third assignment, GenAI feedback. The researcher scored each assignment both before and after feedback based on 20 rubric elements. The researcher entered the data into SPSS using eight variables: participant number, element number of the rubric, Assignment 1 pre-score, Assignment 1 post-score, Assignment 2 pre-score, Assignment 2 post-score, Assignment 3 pre-score, and Assignment 3 post-score. This resulted in a total of 320 rows of data, as each of the 16 participants had scores for 20 rubric elements. Using the Select Cases option in SPSS, the researcher systematically selected each rubric element and calculated the mean scores for each element and assignment by using the descriptive statistics function to calculate and report the mean scores before and after feedback for each assignment. This approach helped determine the average improvements for each rubric element resulting from the different types of feedback.

3.6.3. Analysis Procedures for Research Question 3

Regarding the Research Question 3, the study was intended to shed light on the perceived strengths and weaknesses of feedback provided by teacher, peers, and GenAI. In order to do this, six questions were asked to the participants during the individual interviews. The interviews were conducted via Zoom and saved. Then, the analysis procedures started and expanded to three levels of coding as stated by Dörnyei (2007).

Dörnyei (2007) introduced the procedures of qualitative content analysis including pre-coding, initial coding, second-level coding, display of the ideas, interpretation of the data and drawing conclusions. However, Dörnyei (2007) also introduced a similar qualitative content analysis for grounded theory consisting of three levels of coding including open coding, axial coding, and selective coding. For the analysis, blend of qualitative content analysis procedures and grounded theory coding levels were employed according to Dörnyei (2007).

Firstly, the researcher fully transcribed the interviews. Then, as the interviews were conducted on the participants' native language, they were translated into English by the researcher. After the translation, the translated interview texts and original transcriptions in their native language were sent to the participants individually for them to check whether there were any different meanings occurred during the translation. After the participants approved their translations, the initial coding stage was conducted.

In the pre-coding stage, the researcher used a highlighter and read all the transcriptions by highlighting any sentence, statement, or phrase on the soft copies of the documents that might be beneficial for answering the Research Question 3. This stage was taken from Dörnyei's (2007) qualitative content analysis. In the second stage, using the open coding from coding in grounded theory (Dörnyei, 2007), the highlighted chunks were assigned labels by questioning what each chunk exemplified. In the third stage, axial coding (Dörnyei, 2007) were applied to the labels by collecting them into broader categories. In the fourth stage, selective coding (Dörnyei, 2007) included the selection of a core category

covering all the sub-categories formed before. In the fifth stage, the display of the ideas from Dörnyei's (2007) qualitative content analysis was used by the researcher by presenting interview profiles of the participants in narration and complying the themes from the interview profiles. For the stages included in the qualitative content analysis were presented in the Figure 9.

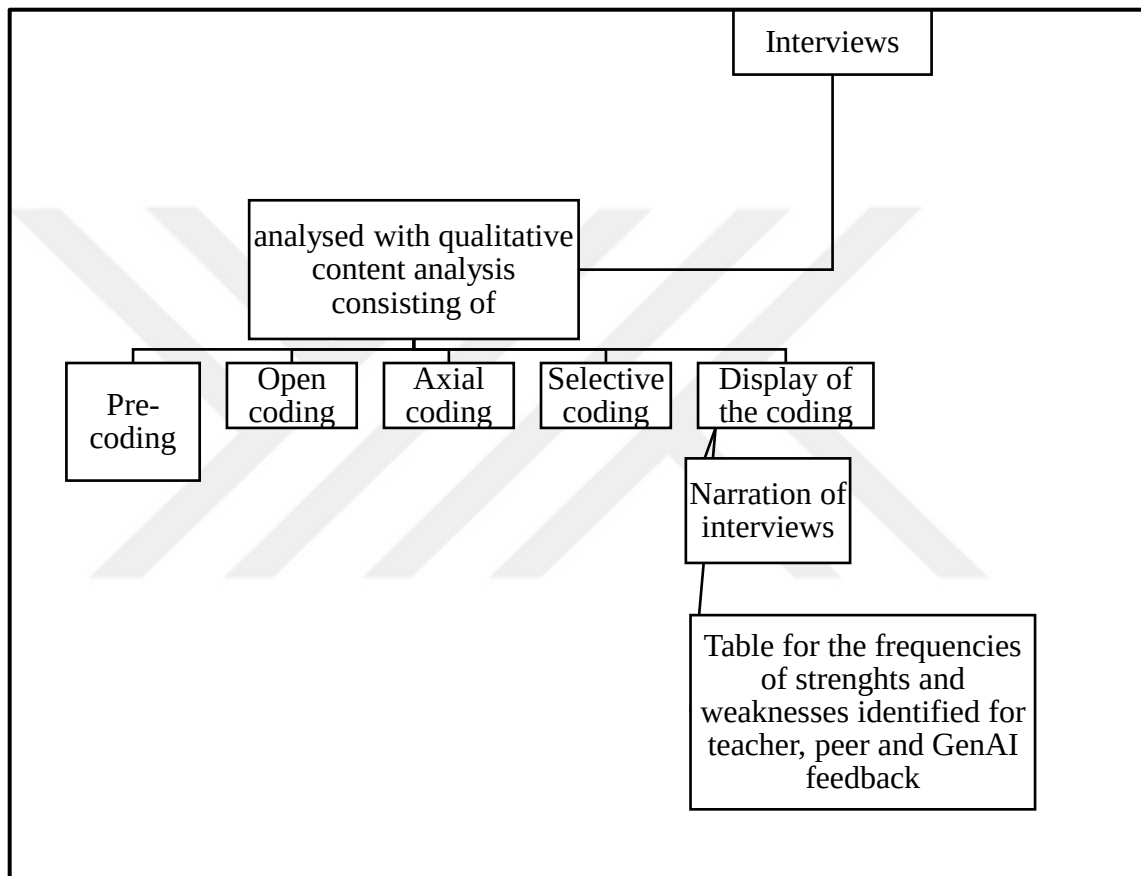


Figure 9. Data Analysis Procedures for Research Question 3

3.6.4. Inter-rater Reliability

In order to establish inter-rater reliability for the assessment of student papers, an independent rater who works as an EFL teacher was invited to rate a selection of papers. For this purpose, the independent rater was familiarized with the rubric and trained how to score papers with reference to the 20 items in the rubric. Then, the rater was given a total number of 30 papers, 5 from each of the three assignments representing both pre and post-tests. The independent rater scored all those 30 papers by using the rubric and the scores coming from

the independent rater were compared to the scores of the researcher of the present study by means of Kendall's tau-b correlation. The results indicated a strong correlation ($r = .83$, $p < .001$) for the pre test scores whereas they revealed a strong correlation ($r = .84$, $p < .001$) for the post test scores.

3.7. The Summary of the Chapter

Chapter Three outlined the comprehensive methodology employed in this study, focusing on research design, setting and participants, data collection tools and sources, data collection procedures, and data analysis methods. The study used a mixed method approach within an instrumental collective case study design. The primary aim was to compare the effectiveness of feedback provided by teachers, peers, and GenAI on students' essay scores and to analyse the strengths and weaknesses of these feedback providers. By integrating both qualitative and quantitative data, the research aimed to provide a holistic understanding of the effects of different feedback types.

The study addressed three main research questions: the differences in essay score improvements based on feedback from teachers, peers, and ChatGPT; the impact of these feedback sources on the improvement of students' assignment grades; and students' perceptions of the strengths and weaknesses of feedback from teachers, peers, and ChatGPT. The study was conducted with 16 students from the Writing Skills I course at the Department of English Language Teaching. The course structure and participant selection process were detailed, emphasizing the authentic and organic setting of the study. Participants were selected based on their scores, ensuring a representative sample. Data collection involved various tools and sources: the Fundamentals of Academic Writing Rubric (FAWR) by Razi (2023) for evaluating assignments, three essay assignments as primary data sources, GenAI prompts developed for obtaining feedback from ChatGPT, and interview questions designed to gather qualitative insights from participants. Detailed procedures for data collection included pre- and post-scores for Assignment 1 with teacher feedback provided via Turnitin, pre-scores and three peer feedback datasets for Assignment 2, and pre-scores, ChatGPT feedback scores, and post-scores for Assignment 3. Participant interviews were conducted in Turkish, transcribed, translated, and verified by participants. Data analysis involved both

qualitative and quantitative methods. For Research Question 1, the Wilcoxon signed ranks test was used to compare gain scores and pre-scores from different feedback sources. Research Question 2 employed qualitative narration and descriptive statistics to analyse feedback effects on assignment scores. Research Question 3 implemented qualitative content analysis with three levels of coding (open, axial, selective) to interpret interview data. The methodology chapter provided a thorough explanation of the research design, data collection, and analysis procedures, ensuring the study's robustness and reliability.



CHAPTER FOUR

FINDINGS

This mixed-method collective case study aims to ethically implement GenAI in the process of academic writing and investigate the impacts of feedback from the teacher, peers, and GenAI on the essay scores measured by pre- and post-feedback scores of the first-year students at a state university in Türkiye in English Language Teaching Department enrolled in Academic Writing Skills I course. The research questions aim to explore three aspects: the comparison of the effect of feedback provided by the teacher, peers, and GenAI, the comparison of the pre-scores given by the teacher and the scores given by peers and GenAI, and the effect of feedback from the teacher, peers, and GenAI on the scores of the assignments. The third aspect includes the strengths and weaknesses of the feedback from the teacher, peers, and GenAI. The first two aspects are analyzed using quantitative data, while the third aspect involves qualitative data gathered from structured interviews. The findings of the research are detailed in this chapter, which consists of eleven sections and a summary.

4.1. Comparative effect of Feedback Provided by the Teacher, Peers, and Generative Artificial Intelligence

4.1.2. The Difference between the Gain Scores of the Feedback Provided by Teachers, Peers, and Generative Artificial Intelligence

Upon calculating the gain score of each assignment by subtracting the pre-score from the post-score for participants' assignments, Wilcoxon Signed Ranks Test with paired samples was conducted to determine whether the population mean ranks differ. Three comparisons were made: the gain scores of Assignment 2 versus Assignment 1, Assignment 3 versus Assignment 1, and Assignment 3 versus Assignment 2. The results showed significant differences in two comparisons: Assignment 2 versus Assignment 1 and Assignment 3 versus Assignment 1.

On the first direction of the comparison, the gain scores of Assignment 2 versus Assignment 1 were analysed. Regarding negative ranks, there were 12 instances where the

gain score for Assignment 2 was lower than the gain score for Assignment 1, with a mean rank of 8.21 and a sum of ranks of 98.50. Regarding positive ranks, there were 2 instances where the gain score for Assignment 2 was higher than the gain score for Assignment 1, with a mean rank of 3.25 and a sum of ranks of 6.50. There were 2 ties, where the gain scores were equal. The test revealed a statistically significant difference between the gain scores of Assignment 2 and Assignment 1, with Assignment 1 having higher gain scores than Assignment 2.

In the second direction of the comparison, the gain scores of Assignment 3 versus Assignment 1 were analysed. There were 12 instances where the gain score for Assignment 3 was lower than that for Assignment 1, with a mean rank of 7.13 and a sum of ranks of 85.50. Regarding positive ranks, there was 1 instance where the gain score for Assignment 3 was higher than the gain score for Assignment 1, with a mean rank of 5.50 and a sum of ranks of 5.50. There were 3 ties. The test indicated a statistically significant difference between the gain scores of Assignment 3 and Assignment 1, with the gain scores for Assignment 1 being higher than those for Assignment 3.

In the third comparison, the gain scores of Assignment 3 versus Assignment 2 were analysed. Regarding negative ranks, there were 6 instances where the gain score for Assignment 3 was lower than the gain score for Assignment 2, with a mean rank of 6.58 and a sum of ranks of 39.50. Regarding positive ranks, there were 6 instances where the gain score for Assignment 3 was higher than the gain score for Assignment 2, with a mean rank of 6.42 and a sum of ranks of 38.50. There were 4 ties. The test indicated no statistically significant difference between the gain scores of Assignment 3 and Assignment 2.

In summary, the gain scores of Assignment 1 were generally superior to both Assignment 2 and Assignment 3, indicating that participants showed more improvement from teacher feedback in Assignment 1 compared to subsequent assignments. The Wilcoxon Signed Ranks Test revealed significant differences when comparing the gain scores of Assignment 1 with both Assignment 2 and Assignment 3, emphasizing the effectiveness of teacher feedback. However, no significant difference was found between the gain scores of

Assignment 2 and Assignment 3, suggesting that participants' progress remained relatively consistent between these two assignments.

Table 3

Wilcoxon Signed Ranks Test Results for Comparison of Gain Scores Across Different Feedback Types for Essay Assignments (N=16)

Direction of comparison		<i>N</i>	<i>Mean rank</i>	Sum of ranks	<i>z</i>	<i>p</i>
Gain score of Assignment 2 vs. Assignment 1	Negative ranks	12	8.21	98.50	-2.90	.004
	Positive ranks	2	3.25	6.50		
	Ties	2				
	Total	16				
Gain score of Assignment 3 vs. Assignment 1	Negative ranks	12	7.13	85.50	-2.80	.005
	Positive ranks	1	5.50	5.50		
	Ties	3				
	Total	16				
Gain score of Assignment 3 vs. Assignment 2	Negative ranks	6	6.58	39.50	-.04	.969
	Positive ranks	6	6.42	38.50		
	Ties	4				
	Total	16				

4.1.3. The Difference between the Pre-scores given by the Teacher and the Scores Given by Peers and Generative Artificial Intelligence

To determine whether there was a significant difference between the pre-scores given by the teacher and the scores given by the peers and GenAI, Wilcoxon Signed Ranks Tests with paired samples were conducted for the Assignment 2 and Assignment 3.

On the first direction of the comparison, mean value of the peer scores for the Assignment 2 and the pre-scores of the teacher for the Assignment 2 were compared. Regarding negative ranks, there were four instances where the peer score was lower than the teacher's pre-score with the mean rank of 3.50 and the sum of ranks of 14.00. Regarding the

positive ranks, there were 12 instances where the peer score was higher than the teacher's pre-score with the mean rank of 10.17 and sum of ranks of 122.00. Regarding the ties, no ties were encountered. As a conclusion, the test indicated a statistically significant difference between the mean value of the peer scores for the Assignment 2 and the pre-scores of the teacher for the Assignment 2 with peer scores being higher than the teacher pre-scores.

On the second direction of comparison, GenAI score for the Assignment 3 and the pre-scores of the teacher for the Assignment 3 were compared. Regarding the negative ranks, there were seven instances where the GenAI score was lower than the teacher's pre-score with the mean rank of 8.00 and sum of ranks of 56.00. Regarding the positive ranks, there were nine instances where the GenAI score was higher than the teacher's pre-score with the mean rank of 8.89 and sum of ranks of 80.00. Regarding the ties, no ties were encountered. As a conclusion, the test indicated no statistically significant difference between the GenAI score for the Assignment 3 and the pre-scores of the teacher for the Assignment 3.

In summary, there was a significant difference between the pre-scores given by the teacher and the scores given by the peers for the Assignment 2, with peer scores being higher. However, there is no significant difference between the pre-scores given by the teacher and the scores given by GenAI for the Assignment 3.

Table 4

Wilcoxon Signed Ranks Test Results for Comparison of Pre-scores across Different Feedback Types for Essay Assignments (N=16)

Direction of comparison		<i>N</i>	<i>Mean rank</i>	Sum of ranks	<i>z</i>	<i>p</i>
Assignment 2 peer score mean vs. Assignment 2 pre-score of the teacher	Negative ranks	4	3.50	14.00	-2.79	.005
	Positive ranks	12	10.17	122.00		
	Ties	0				
	Total	16				
Assignment 3 GenAI score vs. Assignment 3 pre-score of the teacher	Negative ranks	7	8.00	56.00	-.62	.53
	Positive ranks	9	8.89	80.00		
	Ties	0				
	Total	16				

4.2. The Effects of Feedback from the Teacher, Peers, and Generative Artificial Intelligence on the Scores of the Assignments

In this section, a detailed explanations of the scores received by each participant for each assignment categorised under the feedback provider was presented.

4.2.1. The Effect of Teacher Feedback

P1, on her first assignment, was initially scored as 67.50 before receiving feedback from the teacher. After being presented with rubric feedback and scoring, P1 showed improvement in several aspects of the essay, including the formality level, complexity, flow of ideas, unity, coherence, and concluding remarks. However, citations (paraphrases and summaries), direct quotations, abbreviations, and paragraph development remained unrevised. A 12.5-point increase in total was observed, suggesting that P1 made significant improvement efforts based on the teacher's feedback.

P2, on her first assignment, was initially scored as 77.50 before receiving feedback from the teacher. After being presented with rubric feedback and scoring, P2 demonstrated a substantial improvement in her post-score, achieving a perfect score of 100. Elements such as citations, citation needs, abbreviations, paragraph development, flow of ideas including unity and coherence, and concluding remarks, which were initially scored as 2.50 or zero, were improved to five points after receiving rubric explanations. Consequently, P2 achieved a 22.50-point increase in her score as a result of the teacher's feedback.

P3, on his first assignment, was initially scored as 40 before receiving feedback from the teacher. After being presented with rubric feedback and scoring, P3 showed an increase in several elements. Elements such as topic relevance, topic introduction, abbreviations, grammar, spelling, punctuation, and short forms were scored as five. Passive voice and complexity were scored as 2.50, while the rest of the elements, including thesis statement, citations, direct quotations, citation needed, formality level, structure of the essay, main idea and thesis statement relevance, conclusions and main idea relevance, paragraph development, flow of ideas including unity and coherence, and concluding remarks, were scored as zero. After receiving rubric scorings and explanations, a 7.50-point increase was observed in the post-score. While the scores of the thesis statement, formality level, and passive voice increased by 2.50 points each, the rest of the elements either remained the same or increased to 2.50. Despite this, a 7.50-point increase was achieved in the first assignment of P3 upon teacher feedback.

P4, on her first assignment, was initially scored as 60 before receiving feedback from the teacher. After being presented with rubric feedback and scoring, P4 received scores of 2.50 for the elements including thesis statement, citations, formality level, complexity, structure of the essay, main idea and thesis statement relevance, conclusions and main idea relevance, and paragraph development. Elements such as citation needed, abbreviations, flow of ideas including unity and coherence, and concluding remarks were scored as zero. The remaining elements, including topic relevance, topic introduction, direct quotations, grammar, spelling, punctuation, and short forms, were scored as five. After receiving rubric scoring and explanations, P4's post-score increased to 95, showing improvements in all

elements initially scored as 2.50 or zero. However, P4 was scored as 2.50 on the elements of citation needed and concluding remarks, which were previously scored as zero. Consequently, P4 achieved a 35-point score increase in the post-score after receiving teacher feedback.

P5, on her first assignment, was initially scored as 100 before receiving feedback from the teacher. After being presented with rubric feedback and scoring, although no increase was observed in P5's post-score, a review of the paper revealed improvements made by P5, indicating an effort to enhance an already strong submission. Thus, P5 was able to confirm her strengths as reflected in the 100-point score awarded by the teacher's feedback.

P6, on her first assignment, was initially scored as 45 before receiving feedback from the teacher. After being presented with rubric feedback and scoring, the elements scored as 2.50 included the thesis statement, formality level, complexity, and the structure of the essay, while elements such as citations, direct quotations, citation needed, abbreviations, main idea and thesis statement relevance, conclusions and main idea relevance, paragraph development, flow of ideas including unity and coherence, and concluding remarks were scored as zero. Upon receiving feedback from the teacher, P6's score increased to 52.50. Notably, the thesis statement score increased from 2.50 to five, and the score for abbreviations rose from zero to five, resulting in a 7.50-point increase.

P7, on his first assignment, was initially scored as 57.50 before receiving feedback from the teacher. After being presented with rubric feedback and scoring, the elements of the thesis statement, formality level, complexity, paragraph development, flow of ideas including unity and coherence, concluding remarks, and punctuation were scored as 2.50, while the elements of citations, direct quotations, abbreviations, main idea and thesis statement relevance, and short forms were scored as zero. After receiving feedback from the teacher, a 17.50-point increase was observed. P7 improved the elements of the thesis statement, citation needed, abbreviations, main idea and thesis statement relevance, punctuation, and short forms. Consequently, P7 increased his score from 57.50 to 75.

P8, on her first assignment, was initially scored as 72.50 before receiving feedback from the teacher. After being presented with rubric feedback and scoring, the elements of the thesis statement, citations, formality level, complexity, structure of the essay, paragraph development, and flow of ideas including unity and coherence were scored as 2.50, while the elements of citation needed, and abbreviations were scored as zero. After receiving feedback scoring from the teacher, there was a 20-point increase observed. The increase included improvements in the elements of the thesis statement, citations, citation needed, abbreviations, paragraph development, and flow of ideas including unity and coherence. Consequently, P8 increased her score from 72.50 to 92.50 by improving elements that scored lower on the pre-score.

P9, on her first assignment, was initially scored as 100 before receiving feedback from the teacher. After being presented with rubric feedback and scoring, she received a perfect score of 100, indicating a high level of understanding and execution of the assignment requirements. While there was no increase in her post-score, upon closer inspection, it was evident that P9 had made improvements to her paper, further enhancing its quality.

P10, on his first assignment, was initially scored as 92.50 before receiving feedback from the teacher. After being presented with rubric feedback and scoring, the feedback indicated the need for improvements in complexity, abbreviations, and flow of ideas including unity and coherence scored as 2.50. However, P10's assignment was improved only in the complexity element, increasing from 92.50 to 95, suggesting that while the paper was already strong, there were areas where refinement was possible.

P11, on her first assignment, was initially scored as 72.50 before receiving feedback from the teacher. After being presented with rubric feedback and scoring, the pre-feedback evaluation identified specific areas for improvement including the elements of citation needed, conclusions and main idea relevance, paragraph development, flow of ideas including unity and coherence, and concluding remarks, each of which were scored as zero. Regarding the elements scored, the element of complexity was rated at 2.50. Following the scoring and corresponding explanations, P11's score increased by 17.50 points, from 72.50

to 90. The score for the citation needed element improved from zero to 2.50, while the complexity element remained constant at 2.50. Additionally, the scores for conclusions and main idea relevance, as well as flow of ideas including unity and coherence, increased from zero to 2.50. The elements of paragraph development and concluding remarks saw significant improvements, both increasing from zero to five. In summary, the score for P11's first assignment improved from 72.50 to 90 following the incorporation of the feedback.

P12, on her first assignment, was initially scored as 45 before receiving feedback from the teacher. After being presented with rubric feedback and scoring, the evaluation included scoring of zero for various aspects, including the thesis statement, citations, direct quotations, necessity of citations, abbreviations, relevance of the main idea and thesis statement, conclusions and main idea relevance, paragraph development, flow of ideas including unity and coherence, and concluding remarks. Conversely, elements such as complexity, essay structure, and punctuation received a score of 2.50. Following the scoring and corresponding explanations, P12's score increased by 50 points, reaching 95. Significant improvements were observed across several criteria, with the thesis statement, citations (paraphrases and summaries), direct quotations, necessity of citations, abbreviations, relevance of the main idea and thesis statement, conclusions and main idea relevance, paragraph development, flow of ideas including unity and coherence, and concluding remarks all increasing from zero to 5. Additionally, the elements of complexity, essay structure, and punctuation improved from 2.50 to 5. Several aspects, including topic relevance, topic introduction, formality level, passive voice, grammar, spelling, and use of short forms, were consistently scored at 5 throughout the assessment process. In summary, P12's score for the first assignment increased significantly from 45 to 95 after integrating the feedback from the teacher.

P13, on her first assignment, was initially scored as 70 before receiving feedback from the teacher. After being presented with rubric feedback and scoring, her score remained unchanged after receiving feedback from the teacher. The scoring criteria included citations, citation needed, complexity, structure of the essay, conclusions and main idea relevance, and

paragraph development, all scored at 2.50. Formality level, flow of ideas including unity and coherence, and concluding remarks were scored as zero.

P14, on her first assignment, was initially scored as 47.50 before receiving feedback from the teacher. After being presented with rubric feedback and scoring, the elements of the thesis statement, formality level, and complexity were rated 2.50, while citations, direct quotations, citation needed, essay structure, main idea and thesis statement relevance, conclusions and main idea relevance, paragraph development, flow of ideas including unity and coherence, and concluding remarks were all scored zero. After receiving the teacher's feedback, P14's score increased by 2.50 points to 50, specifically in the element of citation needed, resulting in an overall increase from 4.50 to 50.

P15, on her first assignment, was initially scored as 65 before receiving feedback from the teacher. After being presented with rubric feedback and scoring, the elements of citations, formality level, complexity, structure of the essay, conclusions and main idea relevance, and paragraph development were scored as 2.50, while citation needed, abbreviations, flow of ideas including unity and coherence, and concluding remarks were scored as zero. After receiving feedback from the teacher, P15's score increased from 65 to 92.50. Post-feedback, the elements of complexity, flow of ideas including unity and coherence, and concluding remarks were scored as 2.50. Thus, P15 exhibited a 27.50-point increase between her pre- and post-feedback scores.

P16, on his first assignment, was initially scored as 35 before receiving feedback from the teacher. After being presented with rubric feedback and scoring, his score increased to 55 after receiving feedback from the teacher, indicating a 20-point improvement. Initially, elements such as thesis statement, citations, direct quotations, citation needed, formality level, abbreviations, structure of the essay, main idea and thesis statement relevance, conclusions and main idea relevance, paragraph development, flow of ideas including unity and coherence, and concluding remarks were scored as zero, while complexity and short forms were scored at 2.50. Following the teacher's feedback, the scores for thesis statement, citations, direct quotations, citation needed, formality level, abbreviations, and main idea

and thesis statement relevance increased to 2.50, while the remaining elements remained at zero. In conclusion, P16 demonstrated a 20-point improvement.

4.2.2. The Effect of Peer Feedback

P1, on the second assignment, had a pre-score of 65 from the teacher's evaluation but did not achieve any increase following peer feedback. It was noted that P1 received feedback from only two peers instead of three. Peer 1 scored the assignment at 87.50, while Peer 2 scored it at 90. Contrary to the pre-score by the teacher, Peer 1 scored elements such as topic introduction, thesis statement, direct quotations, and punctuation as 2.50, while concluding remarks were scored as 2.50, similar to the teacher's score of zero. Peer 3, who did not provide feedback or a score, led to a discrepancy in the evaluation. Peer 1's scoring of passive voice, complexity, punctuation, and short forms as 2.50 was contrary to the teacher's score of five for each element. No increase was observed in P1's second assignment as a result of the peer feedback.

P2, on the second assignment, had a pre-score of 60 from the teacher's evaluation and achieved an increase of 12.50 points after peer feedback. The first peer scored P2's paper at 100, a 40-point difference. The second peer scored P2's paper at 72.50, with elements such as the thesis statement, direct quotations, citation needs, complexity, structure of the essay, main idea and thesis statement relevance, conclusion and main idea relevance, paragraph development, flow of ideas including unity and coherence, concluding remarks, and grammar being scored similarly to the teacher's pre-score. The third peer also scored P2's paper at 72.50, with seven elements scored similarly to the teacher's pre-score, including complexity, structure of the essay, main idea and thesis statement relevance, conclusion and main idea relevance, paragraph development, flow of ideas including unity and coherence, and concluding remarks. After receiving feedback from three peers, P2 increased her score from 60 to 72.50, achieving a 12.50-point increase.

P3, on the second assignment, had a pre-score of 55 from the teacher's evaluation and achieved an increase of 2.50 points after peer feedback. Pre-revision scores included

2.50 points for elements such as topic introduction, thesis statement, formality level, structure of the essay, conclusions and main idea relevance, and paragraph development, with zero points for elements such as citations, direct quotations, citation needed, complexity, flow of ideas including unity and coherence, and concluding remarks. The first peer scored P3's paper at 90 out of 100, with elements like the thesis statement, formality level, and flow of ideas including unity and coherence receiving scores of 2.50, similar to the teacher's pre-score, while short forms also received 2.50. The second peer scored P3's paper at 25 out of 100, with zero points for topic introduction, thesis statement, citations, direct quotations, citation needed, formality level, paragraph development, flow of ideas including unity and coherence, concluding remarks, and grammar, while other elements were scored as 2.50, apart from topic relevance, which was scored as five. Eight items were scored the same as the teacher's pre-score. The third peer scored P3's paper at 85 out of 100, with elements such as topic introduction, citation needed, abbreviations, and grammar receiving scores of 2.50, and passive voice scored as zero. As a result of the peer feedback, P3 improved his score from 55 to 57.50, with an improvement in the thesis statement element.

P4, on the second assignment, had a pre-score of 57.50 from the teacher's evaluation and achieved an increase of 7.50 points after peer feedback. The first peer scored P4 as 52.5, while the second and third peers scored P4 as 77.5 and 70, respectively. The first peer's scoring matched the teacher's pre-score on eight elements, including thesis statement, citation needed, formality level, structure of the essay, main idea and thesis statement relevance, conclusions and main idea relevance, paragraph development, and concluding remarks. The second peer scored P4's total as 77.50, with scores of 2.50 for topic relevance, direct quotations, abbreviations, main idea and thesis statement relevance, flow of ideas including unity and coherence, grammar, and spelling, while direct quotations and abbreviations were scored as zero. The third peer scored P4 as 65, with scores of 2.50 for thesis statement, citations, direct quotations, citation needed, main idea and thesis statement relevance, paragraph development, flow of ideas including unity and coherence, and grammar, while direct quotations and formality level were scored as zero. As a result of the peer feedback, P4 increased the score from 57.5 to 65 by improving the elements of complexity, conclusions, and main idea relevance, and concluding remarks.

P5, on the second assignment, had a pre-score of 65 from the teacher's evaluation and achieved an increase of 7.50 points after peer feedback. Elements such as citation needed, complexity, structure of the essay, main idea and thesis statement relevance, grammar, spelling, punctuation, and short forms were scored as 2.50, while concluding remarks and conclusions and main idea relevance were scored as zero. Peer feedback varied: the first peer scored P5's work at 87.50, with scores of 2.50 for the structure of the essay, main idea and thesis statement relevance, flow of ideas including unity and coherence, and zero for the thesis statement, with three scores matching the teacher's pre-score. The second peer gave a score of 92.50, with elements like citation needed, paragraph development, and flow of ideas including unity and coherence scored as 2.50, two of which matched the teacher's pre-score. The third peer scored P5's work at 60, with elements such as topic introduction, citations, direct quotations, citation needed, complexity, structure of the essay, main idea and thesis statement relevance, conclusions and main idea relevance, paragraph development, flow of ideas including unity and coherence, grammar, and punctuation scored as 2.50, while thesis statement and abbreviations were scored as zero, with eight scores aligning with the teacher's pre-score. Following the peer feedback, P5's score increased from 65 to 72.50, reflecting improvements in elements such as citation needed, complexity, and short forms, leading to a total increase of 7.50 points.

P6, on the second assignment, had a pre-score of 65 from the teacher's evaluation but did not achieve any increase following peer feedback. Elements such as citation needed, formality level, complexity, structure of the essay, main idea and thesis statement relevance, conclusions and main idea relevance, paragraph development, flow of ideas including unity and coherence, concluding remarks, and short forms were scored as 2.50, while citations and direct quotations were scored as zero. The first peer scored P6's assignment 62.50, with common scores of 2.50 for topic introduction, thesis statement, citations, citation needed, complexity, structure of the essay, main idea and thesis statement relevance, conclusions and main idea relevance, paragraph development, flow of ideas including unity and coherence, concluding remarks, and grammar, and zero for abbreviations. The second peer scored P6's assignment 57.50, with 2.50 for topic introduction, thesis statement, structure of the essay, main idea and thesis statement relevance, conclusions and main idea relevance, flow of ideas including unity and coherence, concluding remarks, grammar, and punctuation, and zero for

citations, passive voice, and complexity. The third peer gave a score of 90, with 2.50 for citations, direct quotations, punctuation, and short forms. Despite receiving feedback from the peers, P6 did not show any increase in her score.

P7, on the second assignment, had a pre-revision score of 32.50 from the teacher's evaluation and achieved an increase of 10 points after peer feedback. Elements including topic introduction, thesis statement, citations, direct quotations, citation needed, complexity, structure of the essay, main idea and thesis statement relevance, conclusions and main idea relevance, flow of ideas including unity and coherence, and concluding remarks were scored as zero, while elements of formality level, grammar, and punctuation were scored as 2.50. The first peer also scored P7 as 32.50, with common scores of zero for thesis statement, citations, direct quotations, citation needed, complexity, main idea and thesis statement relevance, conclusions, and main idea relevance, and concluding remarks, and 2.50 for formality level, passive voice, structure of the essay, paragraph development, flow of ideas including unity and coherence, grammar, and punctuation. The second peer scored P7 as 65, with common scores of 2.50 for thesis statement, citations, direct quotations, formality level, structure of the essay, conclusions and main idea relevance, paragraph development, flow of ideas including unity and coherence, spelling, and short forms, and zero for abbreviations. The third peer scored P7 as 90, with common scores of 2.50 for complexity, structure of the essay, paragraph development, and flow of ideas including unity and coherence. After receiving feedback from three peers, P7's score increased by 10 points, showing improvements in the elements of citations, direct quotations, concluding remarks, and punctuation, raising his score from 32.50 to 42.50.

P8, on the second assignment, had a pre-score of 97.50 from the teacher's evaluation and achieved an increase of 2.50 points after peer feedback. Before receiving feedback, P8 was scored as 97.50, with the element of direct quotations scored as 2.50. After receiving feedback from peers, P8 increased her score from 97.50 to 100. The second peer and the third peer both scored P8 as 100 out of 100. The third peer scored P8 as 87.50 out of 100, with elements such as the thesis statement, main idea and thesis statement relevance, and

conclusions and main idea relevance scored as 2.50, and the topic introduction scored as zero. After receiving feedback from peers, P8 achieved a perfect score of 100.

P9, on the second assignment, had a pre-score of 100 from the teacher's evaluation. Although there was no increase between the pre- and post-scores, the feedback from peers highlighted areas of strength and areas for improvement. The first peer scored P9 as 92.50, with specific elements like the thesis statement, flow of ideas including unity and coherence, and concluding remarks noted for improvement and scored as 2.50. The second peer scored P9 as 95, focusing on elements such as complexity and the relevance of the main idea and thesis statement. The third peer also scored P9 as 92.50, with attention to the relevance of the topic and the absence of abbreviations. However, since P9's pre-score was already 100, it can be concluded that her second assignment was already exemplary and did not require further improvement.

P10, on the second assignment, had a pre-score of 70 from the teacher's evaluation but did not achieve any increase following peer feedback. The instructor's pre-feedback evaluation encompassed various criteria, including the thesis statement, the necessity of citations, formality level, complexity, relevance of the main idea and thesis statement, conclusions, paragraph development, flow of ideas including unity and coherence, concluding remarks, and punctuation, which collectively received a score of 2.50. Notably, the use of direct quotations was scored as zero. Overall, P10's second assignment was awarded a total score of 70 out of 100. Peer one evaluated P10's work with a total score of 70. This assessment included elements such as topic introduction, citations, complexity, relevance of the main idea and thesis statement, grammar, and punctuation, each receiving a score of 2.50. However, the use of direct quotations, the necessity of citations, and abbreviations were all scored as zero. Similarly, the second peer assigned a score of 67.50, evaluating aspects like citations, complexity, relevance of the main idea and thesis statement, conclusions, paragraph development, flow of ideas including unity and coherence, concluding remarks, and grammar with scores of 2.50, while the thesis statement and the necessity of citations were scored as zero. The third peer, assigning a total score of 37.50, rated criteria such as topic relevance, topic introduction, thesis statement, necessity of

citations, passive voice, complexity, essay structure, relevance of the main idea and thesis statement, and flow of ideas including unity and coherence, each with a score of 2.50. In contrast, citations, direct quotations, formality level, abbreviations, conclusions, paragraph development, concluding remarks, and spelling were scored as zero. Despite receiving feedback from peers, P10 did not exhibit any improvement in his second assignment. Consequently, his score remained unchanged at 70 out of 100.

P11, on the second assignment, had a pre-score of 55 from the teacher's evaluation but did not achieve any increase following peer feedback. The initial pre-score assessment awarded 2.50 points each for topic introduction, citation needed, main idea and thesis statement relevance, and spelling. Conversely, elements such as the thesis statement, complexity, conclusions and main idea relevance, paragraph development, flow of ideas including unity and coherence, and concluding remarks were all scored as zero. The first peer feedback rated P11's work with a total score of 75 out of 100. This evaluation assigned 2.50 points each to topic relevance, topic introduction, direct quotations, complexity, main idea and thesis statement relevance, flow of ideas including unity and coherence, grammar, and short forms, while the thesis statement was scored as zero. The second peer provided a higher total score of 90 out of 100, scoring the elements of direct quotations, citation needed, main idea and thesis statement relevance, and punctuation at 2.50 each. The third peer, however, did not provide any scoring or feedback. Despite the feedback from two peers, P11's post-feedback score did not change, remaining at 55 out of 100, identical to the initial pre-score.

P12, on the second assignment, had a pre-score of 90 from the teacher's evaluation and achieved an increase of 2.50 points after peer feedback. This initial evaluation awarded 2.50 points each for the thesis statement, citations, and citation needed elements. Upon receiving feedback from peers, the first peer assessed P12's work and scored it at 80 out of 100. This evaluation included elements such as topic relevance, thesis statement, direct quotations, citation needed, formality level, complexity, relevance of the main idea and thesis statement, and grammar, each receiving a score of 2.50. The second peer provided a higher score of 90, rating elements like citations, necessity of citations, spelling, and

punctuation at 2.50 each. The third peer also awarded P12 a score of 85, with elements such as topic introduction, citations, complexity, conclusions and main idea relevance, grammar, and spelling each receiving a score of 2.50. After receiving feedback from all three peers, P12's score increased from 90 to 92.50. The improvement was specifically in the element of citation, resulting in a 2.50-point increase. In summary, P12's score for the second assignment increased from 90 to 92.50 following the incorporation of feedback from peers, with the main improvement observed in the area of citation.

P13, on the second assignment, had a pre-score of 55 from the teacher's evaluation but did not achieve any increase following peer feedback. The pre-feedback assessment allocated a score of 2.50 to elements such as the thesis statement, necessity of citations, formality level, essay structure, concluding remarks, and grammar. In contrast, elements like citations, direct quotations, complexity, relevance of conclusions and the main idea, paragraph development, and the flow of ideas including unity and coherence, were all scored as zero. Following the feedback from peers, P13's scores varied significantly. The first peer scored P13 at 65, awarding 2.50 points to elements such as topic relevance, citations, necessity of citations, relevance of the main idea and thesis statement, conclusions and main idea relevance, paragraph development, flow of ideas including unity and coherence, concluding remarks, grammar, and spelling, while assigning a score of zero to passive voice and abbreviations. Conversely, the second peer gave P13 a high score of 97.50, focusing solely on the complexity element, which received a score of 2.50. The third peer, aligning closely with the teacher's scoring, awarded P13 a score of 57.50, assigning 2.50 points to elements like the thesis statement, formality level, relevance of the main idea and thesis statement, paragraph development, flow of ideas including unity and coherence, grammar, and punctuation, while scoring elements such as topic relevance, topic introduction, direct quotations, passive voice, and abbreviations as zero. Despite receiving feedback from all three peers, P13's score remained unchanged, showing no improvement in any of the evaluated elements. In summary, P13 did not demonstrate any advancement in his second assignment.

P14, on the second assignment, had a pre-score of 25 from the teacher's evaluation and achieved an increase of 5 points after peer feedback. For the second assignment, P14 was expected to receive feedback scores from three peers, though only two provided feedback. P14's initial score was 25, with zeroes in all elements except topic relevance, abbreviations, grammar, spelling, and punctuation. The first peer awarded P14 a score of 42.50, assigning 2.50 to topic introduction, formality level, passive voice, essay structure, paragraph development, grammar, and punctuation, while giving zeroes to thesis statement, citations, direct quotations, complexity, main idea and thesis statement relevance, conclusions and main idea relevance, flow of ideas including unity and coherence, and concluding remarks. The second peer scored P14 at 60, assigning 2.50 to paragraph development, flow of ideas including unity and coherence, punctuation, and short form, with zeroes in thesis statement, direct quotations, essay structure, main idea and thesis statement relevance, conclusions, and main idea relevance, and concluding remarks. No feedback was received from the third peer. After the feedback, P14's score increased by 5 points to 30, specifically in the element of short form.

P15, on the second assignment, had a pre-score of 40 from the teacher's evaluation but did not achieve any increase following peer feedback. Prior to receiving feedback, P15's initial score was 40 out of 100. The evaluation encompassed various aspects, with scores of 5 assigned to elements such as topic relevance, the need for citations, and issues with abbreviations, grammar, spelling, and short forms. Elements like topic introduction and punctuation received scores of 2.50, while the remaining elements were scored at zero. The first peer's feedback focused on complexity, grammar, and spelling, each scored as 2.50, resulting in an overall score of 92.50 out of 100. The second peer evaluated P15's work, attributing a score of 77.50, which considered aspects such as topic relevance (2.50), thesis statement (2.50), proper citations (2.50), essay structure (2.50), coherence of the main idea and thesis statement (2.50), effectiveness of conclusions (2.50), paragraph development (2.50), concluding remarks (2.50), and punctuation (2.50). The third peer attributed score of zero topic relevance, topic introduction, and the use of direct quotations but assigned a score of 2.50 for grammar. Despite receiving these scores from peers, P15's final score did not increase post-feedback.

P16, on the second assignment, had a pre-score of 57.50 from the teacher's evaluation and achieved an increase of 2.50 points after peer feedback. Elements such as thesis statement, citations (paraphrases and summaries), direct quotations, complexity, structure of the essay, citations needed, and main idea and thesis statement relevance were scored at 2.5, while the rest scored at 0. Peer 1 evaluated P16, giving scores below 5 for elements including topic introduction, citation needed, direct quotations, structure of the essay, paragraph development, grammar, and punctuation, with a total score of 82.50. Peer 2 scored P16 at 70, with most elements scoring 5, except for a few elements including topic introduction, citation needed, direct quotations, structure of the essay, paragraph development, grammar, and punctuation scoring at 2.50. Peer 3 scored P16 at 80, with elements such as topic introduction, complexity, abbreviations, structure of the essay, flow of ideas including unity and coherence, and punctuation scored at 2.50. After receiving peer feedback, P16 increased his score from 57.50 to 60, with an increase in direct quotations from zero to 2.5.

4.2.3. The Effect of Generative Artificial Intelligence Feedback

P1, on the third assignment, was pre-scored as 70 by the teacher and 60 by GenAI. After being exposed to GenAI feedback, a 5-point increase was observed. Common scorings by GenAI and the teacher included the thesis statement, citations, direct quotations, citation needed, formality level, complexity, paragraph development, flow of ideas including unity and coherence, grammar, and punctuation. Contrary to the teacher's pre-scores of five, GenAI scored topic relevance, passive voice, concluding remarks, spelling, and short forms as 2.50. After receiving feedback from GenAI, P1 made several revisions, which resulted in a 5-point increase in her score. These revisions led to improvements in grammar and punctuation. In conclusion, P1 demonstrated a 5-point increase following the feedback and scoring from GenAI.

P2, on the third assignment, was pre-scored as 80 by the teacher and 75 by GenAI. After being exposed to GenAI feedback, there was no change in the score. When the GenAI scoring feedback was inspected, a 5-point difference was observed between the teacher's pre-score and the GenAI score. While the topic introduction, thesis statement, paragraph

development, and grammar were scored the same, concluding remarks were scored as zero by the teacher and 2.50 by GenAI. Similarly, the elements of complexity and flow of ideas including unity and coherence were scored as 2.50 by the teacher and five by GenAI, while the element of passive voice was scored as five by the teacher and 2.50 by GenAI. Despite the scoring and feedback from GenAI, P2 did not make any progress on her third assignment, remaining at a score of 80.

P3, on the third assignment, was pre-scored as 77.50 by the teacher and 55 by GenAI. After being exposed to GenAI feedback, a 17.50-point increase was observed. Initially scored as 77.50 out of 100 by the teacher, elements such as the thesis statement, citations, direct quotations, citation needed, conclusions and main idea relevance, paragraph development, concluding remarks, grammar, and punctuation were scored as 2.50, while the rest were scored as five. The pre-score by the teacher was 22.50 points higher than the GenAI score. GenAI scored direct quotations, structure of the essay, and main idea and thesis statement relevance as five points, while short forms were scored as zero, and the rest of the elements were scored as 2.50. Considering the feedback from GenAI, P3 achieved a total increase of 17.50 points. The same scorings of GenAI and the teacher's pre-score included elements such as the thesis statement, citations, citation needed, conclusions and main idea relevance, paragraph development, grammar, and punctuation. After receiving feedback from GenAI, apart from conclusions and main idea relevance and concluding remarks, which were scored as 2.50, all elements increased to five points.

P4, on the third assignment, was pre-scored as 72.50 by the teacher and 87.50 by GenAI. After being exposed to GenAI feedback, a 2.50-point increase was observed. The pre-score of 72.5 from the teacher include elements such as thesis statement, citations, direct quotations, citation needed, formality level, structure of the essay, main idea and thesis statement relevance, conclusions and main idea relevance, paragraph development, and flow of ideas including unity and coherence scored as 2.50. GenAI scored P4 as 87.50, with elements such as thesis statement, citation needed, formality level, concluding remarks, and grammar scored as 2.50. After receiving feedback from GenAI, P4 showed an increase of 2.50 points in one element of the rubric: formality level.

P5, on the third assignment, was pre-scored as 77.50 by the teacher and 92.50 by GenAI. After being exposed to GenAI feedback, there was no improvement in the score. The pre-score was 77.50, with elements such as complexity, conclusions and main idea relevance, paragraph development, flow of ideas including unity and coherence, concluding remarks, grammar, and punctuation scored as 2.50, and citation needed scored as zero. GenAI feedback scored P5's work at 92.50, with three elements, also common with the teacher's pre-score, scored as 2.50: citation needed, grammar, and punctuation. Despite the feedback, P5 did not improve any elements scored as 2.50, resulting in a final score of 77.50, identical to the pre-score.

P6, on the third assignment, was pre-scored as 55 by the teacher and 62.50 by GenAI. After being exposed to GenAI feedback, there was no change in the score. The teacher's pre-score was 55, with 14 elements scored as 2.50, including topic introduction, thesis statement, citations, formality level, complexity, structure of the essay, conclusions and main idea relevance, paragraph development, flow of ideas including unity and coherence, concluding remarks, grammar, spelling, punctuation, and short forms, while direct quotations and citation needed were scored as zero. GenAI's scoring differed by 7.50 points, with direct quotations and citation needed scored as 2.50 and concluding remarks scored as five. However, despite the feedback from GenAI, P6's score remained unchanged at 55.

P7, on the third assignment, was pre-scored as 60 by the teacher and 72.50 by GenAI. After being exposed to GenAI feedback, there was no change in the score. Elements scored as 2.50 included all except for topic relevance, passive voice, and main idea and thesis statement relevance, which received five points. Compared to the score by GenAI, there was a 12.50-point difference. GenAI scored elements of topic introduction, thesis statement, citations, citation needed, complexity, abbreviations, conclusions and main idea relevance, flow of ideas including unity and coherence, spelling, punctuation, and short forms as 2.50, while the rest were scored as five. Despite the feedback from GenAI, P7 did not improve his assignment, resulting in a stable score of 60.

P8, on the third assignment, was pre-scored as 87.50 by the teacher and 77.50 by GenAI. After being exposed to GenAI feedback, a 7.50-point increase was observed. The pre-score of P8 included elements such as complexity, paragraph development, grammar, spelling, and punctuation scored as 2.50. GenAI scored P8 as 77.50 out of 100, with elements such as citations, direct quotations, passive voice, complexity, paragraph development, concluding remarks, grammar, spelling, and punctuation scored as 2.50. After receiving feedback from GenAI, P8 increased her score from 87.50 to 95. However, the elements of complexity and paragraph development stayed the same as the pre-score, which were scored as 2.50 previously.

P9, on the third assignment, was pre-scored as 97.50 by the teacher and 95 by GenAI. After being exposed to GenAI feedback, there was no change in the score. Before receiving feedback, the teacher's pre-score was 97.50, with the thesis statement identified as an area for improvement. After receiving feedback from GenAI, P9's score remained the same as the pre-score, indicating that while her assignment was strong, there were still some areas where refinement was possible.

P10, on the third assignment, was pre-scored as 87.50 by the teacher and 72.50 by GenAI. After being exposed to GenAI feedback, there was no change in the score. Initially, the pre-feedback score by the teacher was 87.50 out of 100. Five specific elements including topic introduction, thesis statement, necessity of citations, formality level, and complexity were each scored at 87.50. In contrast, GenAI scored P10's third assignment at a total score of 72.50 out of 100. The elements rated at 2.50 included topic introduction, citations, necessity of citations, essay structure, relevance of the main idea and thesis statement, conclusions, paragraph development, grammar, spelling, punctuation, and the use of abbreviations. Despite receiving detailed feedback from GenAI, P10 did not exhibit any improvement in his third assignment compared to the pre-score. Consequently, his post-feedback score remained unchanged at 87.50, the same as the initial pre-feedback evaluation.

P11, on the third assignment, was pre-scored as 67.50 by the teacher and 80 by GenAI. After being exposed to GenAI feedback, a 7.50-point increase was observed. Initially, the pre-feedback score for P11 was 67.50 out of 100. Specific elements such as topic introduction, citation needed, complexity, structure of the essay, relevance of the main idea and thesis statement, conclusions and main idea relevance, paragraph development, flow of ideas including unity and coherence, and concluding remarks were each scored at 2.50, while the thesis statement and formality level were scored as zero. Following the feedback from GenAI, P11's score improved to 80 out of 100. GenAI scored elements like topic introduction, thesis statement, citation needed, passive voice, and complexity at 80. After incorporating the feedback, P11's score increased from 67.50 to 75. The improvements included an increase in the score for topic introduction from 2.50 to five, formality level from zero to 2.50, and complexity from 2.50 to five. Consequently, there was an increase of 7.50 points in P11's score for the third assignment as a result of the feedback from GenAI.

P12, on the third assignment, was pre-scored as 77.50 by the teacher and 72.50 by GenAI. After being exposed to GenAI feedback, a 10-point increase was observed. The pre-feedback assessment, before receiving feedback from GenAI, assigned scores of 2.50 to elements such as the thesis statement, direct quotations, flow of ideas including unity and coherence, grammar, and spelling, while citation needed, and formality level were scored as zero. GenAI's assessment closely aligned with the teacher's evaluation, scoring P12 at 72.50. This evaluation encompassed elements such as citations, direct quotations, citation needed, formality level, passive voice, essay structure, paragraph development, flow of ideas including unity and coherence, grammar, spelling, and punctuation, all scored at 2.50. Following the feedback and scoring from GenAI, P12 improved his score from 77.50 to 87.50. This 10-point increase was attributed to enhancements in the formality level, which increased from zero to 2.50, flow of ideas including unity and coherence, which increased from 2.50 to five, grammar, which improved from 2.50 to five, spelling, which rose from 2.50 to five, and punctuation, which increased from 2.50 to five. In summary, P12's score on the third assignment increased from 77.50 to 87.50 after incorporating feedback from GenAI.

P13, on the third assignment, was pre-scored as 47.50 by the teacher and 75 by GenAI. After being exposed to GenAI feedback, there was no improvement in the score. Initially, the pre-feedback score from the teacher for P13's third assignment was 47.50 out of 100. The evaluation identified specific areas for improvement: topic introduction, necessity of citations, formality level, paragraph development, and concluding remarks, each scored at 2.50, while elements such as the thesis statement, citations, direct quotations, complexity, essay structure, relevance of the main idea and thesis statement, conclusions and main idea relevance, and flow of ideas including unity and coherence were all scored as zero. GenAI's assessment of P13's third assignment resulted in a score of 75 out of 100. This evaluation assigned 2.50 points each to elements such as topic introduction, citations, direct quotations, formality level, passive voice, concluding remarks, grammar, spelling, punctuation, and use of short forms. Despite receiving feedback from GenAI, P13 did not demonstrate any improvement in the scoring of the paper. Consequently, the score for P13's third assignment remained the same. In summary, P13's score for the third assignment did not improve after receiving feedback from GenAI.

P14, on the third assignment, was pre-scored as 32.50 by the teacher and 75 by GenAI. After being exposed to GenAI feedback, there was no change in the score. Initially, the teacher scored P14 at 32.50, assigning zero to topic relevance, thesis statement, citations, direct quotations, citation needed, essay structure, main idea and thesis statement relevance, conclusions and main idea relevance, flow of ideas including unity and coherence, and concluding remarks, and 2.50 to formality level, passive voice, complexity, spelling, and punctuation. ChatGPT scored P14 at 75, assigning 2.50 to citation needed, formality level, passive voice, complexity, flow of ideas including unity and coherence, and grammar, while citations and direct quotations were scored zero. Despite the feedback, P14's score for the third assignment remained at 32.50.

P15, on the third assignment, was pre-scored as 77.50 by the teacher and 80 by GenAI. After being exposed to GenAI feedback, a 2.50-point increase was observed. Initially, the teacher scored P15 as 77.50, assigning 2.50 to complexity, structure of the essay, conclusions and main idea relevance, paragraph development, flow of ideas including

unity and coherence, concluding remarks, and punctuation, while citation needed was scored as zero. ChatGPT scored P15 as 80, assigning 2.50 to citations, citation needed, formality level, structure of the essay, concluding remarks, grammar, spelling, and punctuation. After receiving feedback from ChatGPT, P15's score increased to 80. This increase was due to the element of citation needed, which rose from zero to 2.50 points. Thus, a 2.50-point increase was observed in P15's third assignment.

P16, on the third assignment, was pre-scored as 80 by the teacher and 60 by GenAI. After being exposed to GenAI feedback, there was no change in the score. Initially, P16 scored 80 out of 100. The pre-feedback scores included 5 for topic relevance, topic introduction, thesis statement, citations (paraphrases and summaries), direct quotations, formality level, passive voice, abbreviations, main idea and thesis statement relevance, conclusions and main idea relevance, spelling, punctuation, and short forms; 2.50 for complexity, structure of the essay, paragraph development, flow of ideas including unity and coherence, concluding remarks, and grammar; and zero for citation needed. ChatGPT scored P16 as 60, assigning 2.50 to topic introduction, thesis statement, citations (paraphrases and summaries), direct quotations, citation needed, formality level, passive voice, conclusions and main idea relevance, paragraph development, flow of ideas including unity and coherence, concluding remarks, grammar, spelling, punctuation, and short forms, and 5 to topic relevance, abbreviations, structure of the essay, and main idea and thesis statement relevance. Despite the feedback from GenAI, P16's post-feedback score remained at 80. In conclusion, P16 demonstrated improvements in various elements across the assignments, particularly in response to teacher and peer feedback. Despite receiving feedback from GenAI, P16's score did not show improvement.

4.2.4 The Impact of Feedback on Rubric Elements across Assignments

After examining the improvement in scores for each participant, each rubric element was analysed to determine the mean score. Based on this analysis, Table 5 was created to present the pre- and post-scores for 20 rubric elements across three assignments. In light of

the analysis conducted, the findings were presented to illustrate how each type of feedback uniquely contributed to improving different aspects of the assignments.

Table 5

Pre- and Post-Score Means for 20 Rubric Elements Across Three Assignments

Element no	Assignment 1		Assignment 2		Assignment 3	
	Pre-score mean	Post-score mean	Pre-score mean	Post-score mean	Pre-score mean	Post-score mean
1	5.00	5.00	5.00	5.00	5.00	5.00
2	5.00	5.00	3.75	3.75	3.75	3.91
3	3.13	4.38	2.97	3.13	2.66	2.81
4	2.03	3.13	2.34	2.66	3.44	3.59
5	2.81	3.28	2.03	2.50	3.13	3.28
6	1.88	3.59	2.34	2.66	2.03	2.34
7	2.97	3.91	3.13	3.13	3.28	3.75
8	4.84	5.00	4.38	4.38	4.84	4.84
9	2.97	3.44	1.56	1.88	2.97	3.13
10	2.03	4.22	5.00	5.00	5.00	5.00
11	3.13	3.59	2.34	2.50	3.44	3.44
12	2.97	3.75	2.66	2.81	4.06	4.06
13	2.66	3.44	1.88	2.03	3.28	3.28
14	2.19	3.28	1.72	1.72	2.81	2.97
15	1.09	2.66	1.72	1.88	2.81	2.97
16	1.56	2.97	1.72	2.19	2.97	2.97
17	5.00	5.00	4.22	4.22	3.59	4.22
18	5.00	5.00	4.69	4.69	4.22	4.53
19	4.69	5.00	4.06	4.22	3.75	4.22
20	4.53	5.00	4.38	4.84	4.69	4.69
Total mean	3.27	4.03	3.09	3.26	3.59	3.75

Regarding the teacher feedback, a huge impact on various rubric elements across the assignments was observed. For the thesis statement (3), there was an increase of more than one point from 3.13 to 4.38. For citations (4), the mean score increased from 2.03 to 3.13. Direct quotations (5) improved from 2.81 to 3.28 in Assignment 1. The element of citation needed (6) rose from 1.88 to 3.59 in Assignment 1. Additionally, formality level (7)

increased from 2.97 to 3.91 in Assignment 1. Complexity (9) improved from 2.97 to 3.44 in Assignment 1. Grammar (17) significantly improved from 2.03 to 4.22 in Assignment 1. Some elements experienced slight increases, such as the structure of the essay (11), which improved from 3.13 to 3.59, and the main idea and thesis statement relevance (12), which increased from 2.97 to 3.75. Conclusions and main idea relevance (13) saw an increase from 2.66 to 3.44, while paragraph development (14) improved from 2.19 to 3.28. The flow of ideas including unity and coherence (15) showed a rise from 1.09 to 2.66 and concluding remarks (16) increased from 1.56 to 2.97.

Regarding peer feedback, the analysis of the second assignment revealed notable improvements in several rubric elements. For the thesis statement (3), there was a slight increase from 2.97 to 3.13. Citations (4) improved from 2.34 to 2.66. Similarly, direct quotations (5) saw an increase from 2.03 to 2.50. The element of citation needed (6) also improved from 2.34 to 2.66. Additionally, formality level (7) remained the same at 3.13. Paragraph development (14) did not change, remaining at 1.72. Flow of ideas including unity and coherence (15) increased slightly from 1.72 to 1.88. Concluding remarks (16) improved from 1.72 to 2.19. Some elements experienced slight increases, such as the structure of the essay (11), which improved from 2.34 to 2.50, and the main idea and thesis statement relevance (12), which increased from 2.66 to 2.81. Conclusions and main idea relevance (13) saw an increase from 1.88 to 2.03. However, some elements remained the same, topic introduction (2) at 3.75, passive voice (8) at 4.38, grammar (17) at 4.22, and spelling (18) at 4.69. Punctuation (19) improved slightly from 4.06 to 4.22. The element of short forms (20) increased from 4.38 to 4.84.

Regarding GenAI feedback, the analysis of the third assignment revealed several improvements across various rubric elements. For the thesis statement (3), there was a slight increase from 2.66 to 2.81. Citations (4) improved from 3.44 to 3.59. Similarly, direct quotations (5) saw an increase from 3.13 to 3.28. The element of citation needed (6) improved slightly from 2.03 to 2.34. Additionally, formality level (7) increased from 3.28 to 3.75, indicating better adherence to formal writing standards. Complexity (9) improved slightly from 2.97 to 3.13. Grammar (17) saw a notable improvement from 3.59 to 4.22.

Spelling (18) also improved from 4.22 to 4.53, reflecting consistent accuracy in spelling. Punctuation (19) improved from 3.75 to 4.22. Some elements experienced slight increases, such as paragraph development (14), which improved from 2.81 to 2.97, and the flow of ideas including unity and coherence (15), which increased from 2.81 to 2.97. However, some elements remained the same, including structure of the essay (11) at 3.44, main idea and thesis statement relevance (12) at 4.06, conclusions and main idea relevance (13) at 3.28, concluding remarks (16) at 2.97, passive voice (8) at 4.84, and short forms (20) at 4.69.

4.3. Strengths and Weaknesses of Feedback from the Teacher, Peers, and Generative Artificial Intelligence

In this section, the responses to six questions asked during interviews with participants were narrated in detail. This section consists of the section of themes and three sub-sections including each feedback providers' strengths and weaknesses collected from the participants during the interviews. Each feedback providers' strengths and weaknesses were explained in paragraphs as stated by the participants. Participants answered questions regarding the strengths and weaknesses of teacher feedback, peer feedback, and GenAI feedback. Following the narrative of cases based on participants, the evident themes identified through qualitative content analysis were narrated. These themes were then presented in a table, categorizing the strengths and weaknesses of teacher feedback, peer feedback, and GenAI feedback.

4.3.1. Strengths and Weaknesses of Teacher Feedback

On her first assignment, P1 received a score of 67.50 out of 100 from the teacher. When interviewed about the strengths of the teacher's feedback, P1 noted that it was more trustworthy due to the teacher's expertise and familiarity with the course requirements. This expertise enabled the teacher to provide precise feedback that directly contributed to P1's improved assessment scores. Additionally, P1 regarded the teacher as the most accurate source of feedback compared to other providers. However, P1 identified certain weaknesses in the teacher's feedback. One significant issue was the teacher's limited time, attributed to the educational system's assignment of 120 students to the course. This large student-to-

teacher ratio hindered the teacher's ability to offer detailed feedback to each student. Consequently, the feedback provided was often superficial, though still effective. Compared to GenAI, the teacher's feedback was less detailed, potentially overlooking some aspects. In summary, while teacher feedback was evaluated as more trustworthy and accurate by P1, its effectiveness was constrained by time limitations, resulting in less detailed commentary than that provided by GenAI.

On her first assignment, P2 received a score of 77.50 out of 100 from the teacher. During the interview, P2 mentioned that the strength of the teacher's feedback lay in its reliability, thanks to the teacher's expertise. Compared to other feedback providers, teacher feedback was the most accurate as it directly highlighted areas needing improvement. Initially, P2 did not fully understand how to utilize the feedback since it was her first assignment of the term and her first experience receiving feedback. Despite this initial confusion, the feedback proved effective once she understood how to apply it. In conclusion, teacher feedback, though initially unclear, became a valuable resource for improvement once comprehended.

On his first assignment, P3 received a score of 40 out of 100 from the teacher. During the interview, P3 discussed the strengths and weaknesses of the teacher's feedback. He noted that the strengths included the trustworthiness of the teacher, as the feedback provided was not questionable, unlike peer and GenAI feedback. However, P3 noted that the feedback lacked detail, as it was superficially given using a rubric. The absence of specific markings or comments on the paper was also considered a weakness. In conclusion, while teacher feedback was regarded as trustworthy, its superficial nature and lack of detailed comments limited its effectiveness for P3.

On her first assignment, P4 received a score of 60 out of 100 from the teacher. During the interview, P4 emphasized that the most valuable aspect of the teacher's feedback was its source, as the feedback provider was the teacher who had instructed her on writing an academic paper. P4 found this advantageous because it eliminated the need to verify the feedback's accuracy. However, she identified a weakness in the teacher's feedback related

to the scoring methodology. According to P4, the feedback included scoring and its explanation, which led her to fear receiving zero points for elements rated as 2.50, urging her to revise these aspects even when they were already satisfactory. In conclusion, despite some weaknesses about the scoring process, the teacher's feedback was instrumental in significantly improving P4's performance on the first assignment.

On her first assignment, P5 received a score of 95 out of 100 from the teacher. During the interview, she emphasized the strengths of the teacher's feedback, noting its precision and relevance. She appreciated that the feedback was direct and accurately highlighted areas for improvement, which she attributed to the teacher's expertise. After acknowledging this strong aspect, P5 could not identify any weaknesses in the teacher's feedback. Overall, she found the teacher's feedback to be highly effective and beneficial for her learning.

On his first assignment, P7 received a score of 57.50 out of 100 from the teacher. P7 stated that there were no weak points in the teacher's feedback. He described the teacher's feedback as guiding and beneficial, aiding him in improving his academic writing skills. In conclusion, P7 found the teacher's feedback to be highly constructive and instrumental in his development as a writer.

On her first assignment, P8 received a score of 72.50 out of 100 from the teacher. During the interview, P8 highlighted that teacher feedback was the most trustworthy among all feedback sources, citing the teacher's expertise in academic writing as a key factor. She noted that the teacher's feedback effectively identified areas needing improvement, leading to significant enhancements in her essay's academic quality. However, P8 also mentioned a weakness in the teacher's feedback, which was its lack of detail compared to GenAI feedback. She observed that the teacher overlooked some grammar and punctuation errors. Consequently, she expressed a desire for more comprehensive feedback from the teacher to include these aspects. In conclusion, teacher feedback greatly improved her assignment despite its lack of detail in certain areas.

On her first assignment, P9 was awarded a perfect score of 100 out of 100. During the interview, she stated that the feedback from the teacher was highly motivating, as it demonstrated the teacher's commitment to helping her transform a good assignment into an excellent one by providing advanced hints on academic writing. She perceived no weaknesses in the teacher's feedback. However, she mentioned that additional feedback could further enhance her assignment. P9 found the teacher's feedback highly motivating and helpful, with no weaknesses, but wanted more feedback to improve further.

On his first assignment, P10 received a score of 92.50 out of 100 from the teacher. During the interview, he mentioned that the teacher's feedback was helpful for his future academic writing, as he could read and understand the rubric explanations, which were straightforward to apply. However, he pointed out that the feedback lacked detail, as it didn't clearly show where he made mistakes or how to improve specific areas. In conclusion, while the teacher's feedback was generally useful, its lack of detailed critique limited its effectiveness.

On her first assignment, P11 received a score of 72.50 out of 100 from the teacher. She said that having feedback with a rubric helped her identify which aspects needed significant improvement and which required only minor tweaks. The rubric was effective because it used clear scores: zero for "not acceptable," one for "acceptable," and two for "appropriate." However, P11 wished for more detailed feedback, as she initially found it hard to understand what actions were needed. She expected to see annotations on her paper showing which parts needed changes and pointing out specific mistakes. Overall, while the rubric was helpful, P11 felt she needed more explicit guidance to fully understand how to improve.

P12 got a score of 45 out of 100 on his first assignment. He found the teacher's feedback very motivating because it clearly showed his mistakes and areas needing improvement. Initially, P12 thought he had done well, but the feedback made him realize his shortcomings, which motivated him to improve. However, he felt that the feedback, while clear and concise, could have been more detailed and explanatory. Since this was the first

feedback he had received, he didn't fully understand it at first. After discussing it with his peers, he was able to comprehend the feedback and identify areas for improvement. Overall, P12 found the teacher's feedback motivating but initially challenging to understand because of its brevity and his lack of prior experience with feedback.

P13 received a score of 70 out of 100 on his first assignment. He compared getting feedback from the teacher to having a conversation with someone knowledgeable and effective. However, he noted that unlike GenAI, humans can make mistakes and overlook details. He saw the potential for error and imperfection as a weakness of teacher feedback. In summary, while he found the teacher's feedback knowledgeable, he also recognized that it was prone to human error.

P14 got a score of 47.50 out of 100 on her first assignment. She said the teacher's feedback helped her correct various aspects of her essay, such as paragraph structure, rationality, and the formulation of topic and conclusion sentences. However, she expected more comprehensive feedback. She wanted comments that specifically addressed her deficiencies and provided guidance on how to improve them. While the feedback was beneficial, it lacked the depth and specificity she needed for further improvement.

P15 received a score of 65 out of 100 on her first assignment. She appreciated the brevity of the teacher's feedback, which was clear and direct in identifying areas needing improvement. However, she noted that the feedback lacked specificity, leaving her uncertain about the precise issues and their locations. In summary, while she found the clarity of the teacher's feedback beneficial, its lack of detail prevented her from understanding specific areas needing improvement.

P16 got a score of 35 out of 100 on his first assignment. He valued the teacher's feedback and paid close attention to it. The rubric scoring and feedback helped him understand the teacher's scoring system, which he found beneficial as a student. However, since this was his first essay in academic writing, he expected more detailed and explanatory

feedback. In conclusion, while he found the teacher's feedback valuable and helpful for understanding the scoring system, it lacked the detailed guidance he needed for improvement.

4.3.2. Strengths and Weaknesses of Peer Feedback

On the second assignment, P1 received feedback from two peers who scored her as 87.50 and 90. During the interview, P1 stated that the strength of the peer feedback lay in the rigorous evaluation conducted by the peers. According to the course requirements, students were also scored based on the feedback they provided, which motivated peers to spend considerable time giving detailed feedback to secure good scores from the teacher. However, P1 also identified weaknesses in peer feedback. She noted that since peers were at the same learning stage as herself, their feedback was not entirely reliable; peers might have learned incorrectly and could potentially provide misguided advice. Despite this, P1 still found peer feedback to be effective. In conclusion, while peer feedback was detailed and motivated by the course structure, its reliability was compromised by the peers' own learning process.

On the second assignment, P2 received feedback from three peers who scored her as 100, 72.50, and 72.50. When interviewed regarding the strengths and weaknesses of peer feedback, P2 stated that she did not perceive any strong points in the peer feedback since the mistakes identified by her peers were already known to her. Consequently, P2 found peer feedback less effective than other sources. Additionally, she noted a margin of error in peer feedback, as peers were also students and prone to mistakes. In conclusion, peer feedback was regarded as less effective by P2 due to its redundancy and potential for error.

On the second assignment, P3 received feedback from three peers who scored him as 90, 25, and 85. P3 stated that receiving feedback from peers was motivating, as it alleviated the fear of being judged harshly by a teacher. He did not feel ashamed of making mistakes or being evaluated as unsuccessful by peers, in contrast to his fear of being perceived as lazy by the teacher. However, P3 also noted that peer feedback was not detailed and included

some incorrect information, as suggested by the conflicting scores. This inconsistency made it difficult for P3 to determine which feedback to apply to his assignment. In conclusion, peer feedback was motivating and less intimidating for P3, but its lack of detail and inconsistency reduced its overall effectiveness.

On the second assignment, P4 received feedback from three peers who scored her as 52.50, 77.50, and 70. P4 noted that a strong point of peer feedback was that it came from fellow learners who were also continuously developing their academic writing skills. However, she pointed out a major weakness: some peers did not give enough attention to the feedback they provided, resulting in inaccurate or misleading comments. Consequently, P4 felt compelled to double-check the feedback from her peers. In conclusion, while peer feedback had the potential to be beneficial, its inconsistent quality required further examination by P4.

On the second assignment, P5 received feedback from three peers who scored her as 87.50, 92.50, and 60. In the interview, P5 mentioned that the peer feedback allowed her to self-reflect and assess her work, which she considered the only strong point. However, she found the weaknesses of peer feedback to outweigh the strengths. She highlighted the peers' lack of knowledge in academic reading, noting that they were also in the learning stage. This led to her confusion and self-doubt, as she was uncertain whether her peers or she herself was correct. The conflicting scores, such as the 60 points from the third peer, further diminished the utility of the feedback. Ultimately, P5 concluded that peer feedback did not provide any substantial benefits.

On the second assignment, P6 received feedback from three peers, with scores of 62.50, 90, and 57.50. During the interview, P6 highlighted the strengths of peer feedback, noting that peers were good at identifying issues she had overlooked due to the rush to meet the assignment deadline. She appreciated that peers had sufficient time to review the assignment in detail. However, P6 also noted inconsistencies in the thoroughness of the feedback, which led her to question the accuracy and reliability of some of the comments. In conclusion, while P6 found peer feedback valuable for its thoroughness and fresh

perspectives, inconsistencies in the feedback quality raised concerns about its overall reliability for P6.

On the second assignment, P7 received feedback from three peers who scored him as 32.50, 65, and 90. During the interview, he expressed that he did not benefit from peer feedback. He noted that the scores were highly conflicting, with a range from low to high, given by peers who were still learning the fundamentals of academic writing. Furthermore, P7 remarked that peer feedback was often provided negligently, merely to fulfil the assignment requirement. He also mentioned that peers who received low scores from their feedback providers tended to reciprocate by giving similarly poor feedback and low scores. In conclusion, P7 found peer feedback to be inconsistent and unhelpful.

On the second assignment, P8 received feedback from three peers who scored her as 100, 100, and 87.50. In the interview, she expressed dissatisfaction with the peer feedback, stating that it was unhelpful as it failed to identify her mistakes. She observed that the feedback was undetailed, neglecting grammar errors and the flow of ideas, which suggested to her that the peers were not thorough in their evaluations. This led her to believe that peer feedback was provided negligently. In conclusion, peer feedback did not contribute meaningfully to her improvement due to its lack of thoroughness and attention to detail.

On the second assignment, P9 received feedback from three peers who scored her as 92.50, 95, and 92.50. In the interview, she noted that one peer's feedback was as valuable as the teacher's feedback. She appreciated how her peers reminded her of important aspects of academic writing that she had forgotten, highlighting her reliance on peer support to recall class content. However, she also pointed out that some feedback from her peers was incorrect, which she identified as a potential weakness, noting the risk of being misled. Despite this, the overall peer feedback was beneficial in reinforcing her learning.

On the second assignment, P10 received feedback from three peers who scored him as 70, 67.50, and 37.50. During the interview, he mentioned that he could not benefit much

from peer feedback, as peers who were still learning were not more knowledgeable than he was, and they shared the same instructional context. He also suggested that peers could be less knowledgeable, as not all students engaged equally with the instructor. Nonetheless, he valued peer feedback as a friend's opinion, appreciating comments on various aspects of the assignment. However, the conflicting peer scores indicated inconsistency in the feedback. In conclusion, P10 found peer feedback valuable but inconsistent and less beneficial due to the varying levels of peer knowledge from the course.

On the second assignment, P11 received feedback from two peers who scored her as 75 and 90. In the interview, she noted that the first peer merely provided rubric scores, which were not helpful to her, and consequently, she disregarded them. In contrast, the second peer made annotations on the paper, which highlighted some of her knowledge gaps regarding the assignment. Despite this, P11 double-checked each comment, acknowledging that peers were also students and could make mistakes. She remarked that peer feedback was generally superficial, attributing this to the peers' status as students, and concluded that she did not find peer feedback beneficial due to its lack of depth.

On the second assignment, P12 received feedback from three peers who scored him as 80, 85, and 85. During the interview, he stated that the peer feedback was also motivating because it demonstrated that his peers were capable of providing valuable feedback on his assignment. This realization encouraged him to improve and perform as well as his peers. P12 did not identify any weaknesses in the peer feedback. Thus, peer feedback served as a positive reinforcement, motivating P12 to strive for better performance without any weak points.

On the second assignment, P13 received feedback from two peers who scored him as 65 and 97.50. In the interview, he observed that because providing feedback was a course requirement, some students believed they would receive points for their detailed feedback. However, he recognized a weakness in peer feedback when his best friend was assigned to review his work, resulting in overly flattering comments. He suggested this could occur with other students, leading to insincere feedback. Additionally, he mentioned that peers were

neither as knowledgeable as the teacher nor as attentive and detailed as GenAI. In conclusion, peer feedback was found less reliable due to potential bias and lack of expertise by P13.

On the second assignment, P14 received feedback from two peers who scored her as 42.50 and 60. Despite this, P14 acknowledged several strong points in the peer feedback. She noted that peers, who had a better understanding of the assignment's requirements, provided valuable suggestions that helped her correct her mistakes. However, she constantly questioned the validity of the peer feedback, as she was aware that her peers were not experts in academic writing. Consequently, while the peer feedback was helpful, P14 remained sceptical about its accuracy, which limited her trust in its effectiveness.

On the second assignment, P15 received feedback from three peers who scored her as 92.50, 77.50, and 82.50. She explained that due to time constraints during submission, some points in the assignment might have been overlooked. The peer feedback, provided with more time, helped her identify and correct these oversights. Nevertheless, P15 found some peer feedback to be incorrect, though she was able to discern and disregard these inaccuracies. Despite this, the presence of incorrect feedback was a source of frustration. In conclusion, peer feedback was generally helpful in addressing overlooked issues, but its occasional inaccuracies were a source of annoyance for P15.

On the second assignment, P16 received feedback from three peers who scored him as 82.50, 70, and 80. He stated that peer feedback was less trustworthy compared to the teacher's feedback since peers were at the same level as P16. However, receiving feedback from three different peers provided varied perspectives. Despite this, the third feedback was not useful, and P16 mentioned that he did not benefit much from any of the peer feedback. He only identified his mistakes without understanding how to improve them. In summary, while peer feedback offered diverse perspectives, it was generally unhelpful for P16 in terms of actionable improvement.

4.3.3. Strengths and Weaknesses of Generative Artificial Intelligence Feedback

On the third assignment, P1 received feedback from GenAI and was scored 60 by GenAI. In the final part of the interview, P1 was asked about the strengths and weaknesses of GenAI feedback. P1 stated that GenAI provided the most detailed feedback, being available as much and as long as needed. According to P1, traditional feedback providers, such as teachers and peers, are limited by their human capacity and responsibilities to other students. In contrast, GenAI has unlimited capacity and can assist multiple individuals simultaneously, similar to a private tutor. However, P1 also identified several weaknesses in GenAI feedback, including the provision of false information, incorrect guidance, and occasional misunderstandings of the prompt. Despite acknowledging GenAI's unlimited capacity, P1 noted conflicts in her perception, citing instances where GenAI misguided her in academic writing and provided feedback that did not align with course information, such as referencing rules. P1 also mentioned that despite pre-testing the prompts, GenAI sometimes failed to understand them correctly. In conclusion, while GenAI feedback was considered more effective than peer feedback due to its detailed nature and unlimited capacity, its reliability was occasionally compromised by inaccuracies and misunderstandings.

On the third assignment, P2 received feedback from GenAI and was scored 85 out of 100. Regarding the strengths and weaknesses of GenAI feedback, P2 stated during the interview that she did not identify any strong points in the GenAI feedback. She noted a margin of error in GenAI feedback, which made her unsure about its trustworthiness. Nevertheless, P2 acknowledged that she still benefitted from the feedback provided by GenAI. In conclusion, while GenAI feedback had reliability issues, it still offered some benefits to P2.

On the third assignment, P3 received feedback from GenAI and was scored 55 out of 100. P3 stated that GenAI feedback was very time-efficient, as he did not have to wait for feedback like he did with teacher or peer feedback. GenAI provided feedback in seconds, which helped him save time. Furthermore, GenAI could detect small details in the

assignment, such as minor spelling mistakes, which might be overlooked by peers and the teacher. P3 noted that GenAI was particularly good at details. However, he also stated that GenAI was inconsistent in providing feedback on some aspects, such as scoring. P3 tried the same prompt in different chats and received different scores and feedback focusing on different aspects each time. Additionally, P3 encountered instances where GenAI became confused by the prompts, leading him to question its reliability. In conclusion, while GenAI feedback was time-efficient and detailed, its inconsistency and occasional confusion raised concerns about its reliability for P3.

On the third assignment, P4 received feedback from GenAI and was scored 87.50. In her interview, P4 mentioned that GenAI feedback was superior to peer feedback because it provided accurate insights into both the strengths and weaknesses of her assignment. However, she also noted that GenAI tended to assign high scores for various elements, which occasionally misled her. As a result, P4 did not make significant improvements based on GenAI feedback, leading to a lower score from the course instructor compared to her other assignments. In conclusion, while GenAI feedback was more precise, its overly generous scoring sometimes hindered P4's ability to make meaningful improvements.

On the third assignment, P5 received feedback from GenAI and was scored 92.50 out of 100. During the interview, P5 stated that she found no strong points in the GenAI feedback due to encountering different and conflicting responses. She considered GenAI feedback as her own, as she evaluated each piece of feedback for correctness and reliability. Consequently, she felt that it lacked the external perspective she desired and viewed it as a selection of feedback rather than guidance from a more knowledgeable source. Thus, she concluded that GenAI feedback was not beneficial for her learning.

On the third assignment, P6 received feedback from GenAI and was scored 62.50 out of 100. During the interview, P6 acknowledged that the GenAI feedback was highly detailed, providing explanations for each rubric element and pinpointing errors directly in the text. Despite the comprehensive nature of the feedback, P6 expressed doubts about its reliability, given her perception of GenAI as a newly developed tool with potential inaccuracies.

Consequently, she approached the feedback with a degree of scepticism. In conclusion, P6 appreciated the detailed nature of the GenAI feedback but remained cautious about its accuracy due to concerns about the tool's developing nature and reliability.

On the third assignment, P7 received feedback from GenAI and was scored 72.50. During the interview regarding GenAI feedback's strengths and weaknesses, P7 highlighted that he trusted GenAI more than his peers. He stated that GenAI provided correct guidance and effectively showed him how to make improvements. However, he experimented with GenAI by sending separate parts of his assignment and noted an inconsistency: despite not sending the introduction, he received feedback praising it. This made him suspicious of GenAI's feedback. Nonetheless, P7 mentioned that he had received training on such issues from his teacher, recognizing it as a potential error of GenAI. In conclusion, while P7 found GenAI's feedback generally effective, he remained cautious of its occasional inaccuracies.

On the third assignment, P8 received feedback from GenAI and was scored 77.50. She noted that GenAI provided accurate and useful feedback on spelling, the flow of ideas, and referencing, which helped her enhance her assignment. However, she also pointed out that GenAI gave incorrect or misleading feedback on grammar, punctuation, and the use of passive voice. She expressed concern that users unable to distinguish between accurate and inaccurate feedback from GenAI might see a decline in their academic writing quality. In conclusion, while GenAI feedback was beneficial in certain areas, its inaccuracies posed a risk to the overall quality of academic writing for P8.

On the third assignment, P9 received feedback from GenAI and was scored 95 out of 100. She mentioned during the interview that, as a machine, GenAI provided detailed feedback supported by specific sentences or examples from the text, which she found more helpful than peer feedback. However, she also reported that obtaining this detailed feedback required multiple attempts and prolonged interaction with GenAI. Consequently, she expressed doubts about the accuracy of GenAI's feedback due to conflicting responses. Shortly, P9 found GenAI's feedback detailed and helpful but was concerned about its accuracy due to conflicting responses and the time required to obtain useful feedback.

On the third assignment, P10 received feedback from GenAI and was scored 72.50 out of 100. During the interview, P10 stated that he appreciated the comprehensive feedback provided by GenAI, which covered each rubric element in detail and included example sentences from the text needing improvement. However, he noted that GenAI, being a machine, lacked the human touch, expressing feedback in a very sharp and strictly correct manner without the soft-spoken and humane perspective of a human reviewer. In conclusion, while GenAI's feedback was thorough and detailed, its mechanical nature and lack of humane communication were seen as limitations.

On the third assignment, P11 received feedback from GenAI and was scored 80 out of 100. During the interview, she stated that the initial feedback from GenAI was often unrelated, nonsensical, and incorrect. However, after prompting GenAI to rectify its mistakes, it consistently provided corrected and useful feedback. P11 noted that GenAI's feedback was more helpful than peer feedback, as it was more detailed and specific in identifying mistakes and suggesting corrections. In conclusion, while GenAI initially presented some challenges, its feedback ultimately proved to be more beneficial than peer feedback due to its detailed and corrective nature.

On the third assignment, P12 received feedback from GenAI and was scored 72.50 out of 100. During the interview, he noted that the GenAI feedback was exceptionally detailed, offering a deep analysis of his paper. However, P12 expressed concerns about the reliability of the GenAI feedback, as he encountered instances of incorrect evaluations. Consequently, while the detailed analysis from GenAI was appreciated, the accuracy and trustworthiness of the feedback remained a concern for P12.

On the third assignment, P13 received feedback from GenAI and was scored 75 out of 100. He stated that GenAI was beneficial in several ways, being objective as a machine, instantly detecting and showing how to correct mistakes, and being time-efficient since he did not have to wait for peers or the teacher. He also found it less demanding because it provided clear feedback and allowed for further clarification when needed. However, P13 mentioned that GenAI was not entirely trustworthy due to its ongoing development and

evolution. He emphasized that the quality of feedback heavily depended on the prompts given, which could lead to incorrect feedback. Additionally, he felt that GenAI feedback lacked a human touch, making it seem generic and artificial. In summary, while GenAI feedback was efficient and objective, it was considered less trustworthy and lacked humane touch by P13.

On the third assignment, P14 received feedback from GenAI and was scored 75 out of 100. During the interview, P14 revealed that she had not attended a preparatory class before starting her bachelor's program, resulting in a lack of fundamental academic writing skills. However, with the assistance of GenAI, she learned essential aspects of academic writing, such as paragraph development and crafting topic and conclusion sentences. In conjunction with the teacher's feedback, this guidance helped her improve significantly. Nevertheless, P14 pointed out that GenAI's objectivity, while potentially a strength for some, was a weakness for her. She felt that the teacher's familiarity with her development and writing style provided more personalized feedback, whereas GenAI, as an unfamiliar entity, lacked this contextual understanding. Thus, despite its educational value, GenAI's feedback did not help P14 as effectively as the teacher's feedback.

On the third assignment, P15 received feedback from GenAI and was scored 80 out of 100. She appreciated the detailed feedback, which enabled her to focus on specific areas for improvement, such as the essay's structure, which was rated 2.50, indicating that it was good but could be improved. Consequently, she felt that the detailed feedback contributed to her receiving higher scores than expected. However, she also mentioned that GenAI provided overly high scores for certain rubric elements, causing her to overlook some areas needing improvement. Thus, while the detailed feedback was beneficial, the inflated scores from GenAI led to some missed opportunities for further improvement.

On the third assignment, P16 received feedback from GenAI and was scored 60 out of 100. He noted that GenAI's feedback was more detailed and explanatory, helping him question its reliability as it was supported with evidence. Nonetheless, P16 expressed distrust towards GenAI, frequently asking the same prompt to check for consistent feedback. He also

verified the feedback using different sources such as search engines. Although GenAI provided similar feedback consistently, the scores differed. In conclusion, while GenAI's detailed feedback was appreciated, P16's trust issues with the AI led him to seek verification from multiple sources.

4.3.4 Themes Emerged as the Weaknesses and Strengths of Feedback from the Teacher, Peers, and Generative Artificial Intelligence

The interviews were analysed to identify evident themes in the teacher, peer, and GenAI feedback regarding strengths and weaknesses. After the analysis, the evident themes were tabulated with the frequencies of the themes identified under the categories of feedback providers' strengths, and weaknesses in Table 6.

With the analysis of teacher feedback, three themes emerged as the most pronounced strengths of teacher feedback including trustworthiness, brevity, and motivational aspects. Among these, brevity was the most evident, with a frequency of 11, followed closely by trustworthiness, also with a frequency of 11. The motivational aspect was the least evident theme. Trustworthiness stemmed from the teacher's expertise in academic writing. Participants believed that, as the course instructor and a professional in academic writing, the teacher's feedback was unquestionable and required no further verification. The feedback was also regarded as very brief with the frequency of 11, consisting of rubric explanations and corresponding scores. This brevity allowed participants to clearly understand the strengths and weaknesses of their essays without ambiguity, enabling them to take well-guided and direct actions. Additionally, the concise nature of the feedback helped participants understand the teacher's assessment system, as it highlighted precisely where they might lose points. Lastly, with a frequency of two, the motivational aspect of teacher feedback originated from various participant views, including understanding the scoring system, realizing their mistakes, and recognizing the teacher's commitment. Some participants, having received low scores, were motivated to achieve higher ones. Others were driven to correct their mistakes upon realizing them. Furthermore, participants appreciated the time the teacher devoted to providing feedback, which in turn motivated them to improve

themselves. Thus, the analysis underscores the significant impact of brevity and trustworthiness on the effectiveness of teacher feedback, while motivational aspects, though less frequent, also play a role in encouraging student improvement.

Table 6

The Frequencies of the Strengths and Weaknesses Identified for Teacher, Peer, and GenAI Feedback

Feedback provider	Strengths	<i>f</i>	Weaknesses	<i>f</i>
Teacher	Trustworthiness	11	Lack of detail	17
	Brevity	11	Incomprehensibility	4
	Motivational aspect	2	Disappointment	2
Peer	Detailed	7	Undetailed	7
	Motivational aspect	2	Inconsistent	4
			Untrustworthy	10
			Undetailed	7
GenAI			Incorrect	7
	Detailed	10	Mechanical interaction	4
	Reliable	5	Unreliable	18
	High availability and extensive capability	12	Deceptive scoring	2

The weaknesses of teacher feedback included themes of lack of detail, incomprehensibility, and disappointment among participants. The most evident theme was the lack of detail, with a frequency of 17. Participants noted that teacher feedback often lacked sufficient detail, leading to some mistakes being overlooked. This lack of detail was attributed to the teacher's limited time, which prevented the provision of comprehensive feedback to every student enrolled in the course. Consequently, participants perceived the feedback as superficial. Following this, the theme of incomprehensibility was noted with a frequency of four. Participants found the feedback difficult to understand, especially since it was their first feedback and was perceived as superficial. Lastly, participants expressed disappointment with the frequency of two upon receiving the feedback, as they initially

believed they had performed well on their assignments, only to find the feedback indicated otherwise. Overall, the analysis highlights that the primary issues with teacher feedback were its lack of detail and clarity, which led to misunderstandings and disappointment among students.

Two main themes emerged regarding the strengths of peer feedback including its detailed nature with the frequency of seven and its motivational aspects with the frequency of two. Participants noted that peer feedback was detailed due to course requirements, as students needed to provide comprehensive feedback to secure high scores. Peers had more time to give feedback than essay writers, leading to thorough comments, annotations, and highlights on papers. Receiving feedback from three different peers added diverse perspectives, and the feedback often reflected varying levels of understanding among participants. In addition to being detailed, peer feedback was motivational as it lacked the fear of judgement associated with teacher feedback. Recipients felt encouraged by high-quality feedback, perceiving it as a mark of success and aspiring to match their peers' standards. In summary, peer feedback's strengths lie in its detailed nature, driven by course demands, and its motivational impact, stemming from the supportive peer environment.

Regarding the weaknesses of peer feedback, themes of being incorrect, inconsistent, untrustworthy, and undetailed emerged. The theme of being incorrect, with a frequency of seven, originated from peers providing erroneous feedback, which left recipients feeling frustrated and misguided. The theme of being undetailed, also with a frequency of seven, stemmed from peers who gave feedback negligently, merely to fulfil course requirements. Some participants believed that peers who received poor feedback tended to reciprocate with similarly poor feedback. Another weakness was inconsistency with the frequency of four, as receiving feedback from three different peers often led to varying scores and comments, causing confusion among recipients. The most frequently mentioned theme with the frequency of 10 was untrustworthiness, as the issues of incorrectness, inconsistency, and lack of detail led recipients to distrust peer feedback. Additionally, participants felt that peers, being at the same level as them, might be less knowledgeable, further diminishing the perceived value of the feedback. In summary, the weaknesses of peer feedback include

incorrectness, inconsistency, lack of detail, and perceived untrustworthiness, all of which undermine its effectiveness and reliability.

Upon analysing the interviews to identify evident themes regarding the strengths and weaknesses of GenAI feedback, six themes emerged. The strengths included detailed and reliable feedback, high availability, and extensive capability. Conversely, the weaknesses encompassed unreliability, deceptive scoring, and mechanical interaction. The most frequently mentioned strength was the detailed nature of GenAI feedback, with a frequency of ten. Participants noted that it provided the most detailed feedback compared to other sources, identifying even the smallest details in the papers, and offering explanations and suggestions supported by instances from the essays. The second most frequently mentioned strength, with a frequency of six, was GenAI's high availability and extensive capability. Participants appreciated that, unlike other feedback providers, they did not have to wait for GenAI feedback and could engage in further discussions to better understand the feedback. While peers and teachers had limited capacity to provide feedback simultaneously, participants highlighted GenAI's extensive capability to offer feedback on different aspects of the essay. Additionally, GenAI was considered a reliable source of information with a frequency of five, as participants confirmed that its feedback was consistently supported by explanations and instances from their essays. In conclusion, the strengths of GenAI feedback were characterized by its detailed nature, high availability, extensive capability, and reliability.

In terms of weaknesses, the themes of unreliability (frequency of 18), mechanical interaction (frequency of four), and deceptive scoring (frequency of two) emerged. The most frequently mentioned weakness was the unreliability of GenAI feedback, as participants pointed out that, being a tool still under development, GenAI occasionally provided incorrect information. Participants encountered instances where the feedback was contrary to what they had learned in their course, reinforcing their perception of GenAI as an unreliable tool. The second theme was the mechanical nature of interactions with GenAI. Participants described the interaction as artificial and lacking a human touch, preferring the more personalized and knowledgeable engagement with teachers. Finally, the theme of deceptive

scoring emerged. Participants found GenAI's scoring misleading because, although it might evaluate a specific element of the rubric favourably, it did not always reflect this in the overall score, thereby misguiding participants about necessary improvements. In conclusion, GenAI was criticized for its unreliability, mechanical interactions, and potentially deceptive scoring.

4.4. The Summary of the Chapter

This chapter presents the findings of a mixed-method collective case study exploring the impacts of feedback from teachers, peers, and GenAI on the essay scores of first-year students in an Academic Writing Skills I course at a state university in Türkiye. The study investigates three main aspects: the comparative effects of feedback from different sources, the differences between pre-scores given by teachers and scores given by peers and GenAI, and the overall impact of feedback on assignment scores, including strengths and weaknesses.

Quantitative data analysis using the Wilcoxon Signed Ranks Test revealed significant differences in the gain scores between assignments, indicating that teacher feedback was generally more effective than peer and GenAI feedback in improving student performance. No significant difference was found between peer and GenAI feedback, suggesting a relatively consistent impact.

Qualitative data collected via structured individual interviews from the participants provided insights into the perceived strengths and weaknesses of each feedback type. Teacher feedback was praised for its trustworthiness and precision, but it was criticised for being less detailed due to time constraints according to the majority of the participants. Peer feedback was valued for its detailed nature, motivated by course requirements, but was often inconsistent and unreliable due to the peers' own learning stages. GenAI feedback was appreciated for its detailed, time-efficient, and unlimited capacity but faced criticism for occasional inaccuracies and misunderstandings.

The chapter was organized into three main sections detailing the comparative effects of feedback, the differences in pre-scores and post-scores, and a narration of the strengths and weaknesses of each feedback type based on participant responses. The finding of these sections were presented using quantitative data tables and qualitative narratives to provide a comprehensive understanding of the sections included.



CHAPTER FIVE

DISCUSSION, CONCLUSION, AND IMPLICATIONS

In this chapter, the findings of the study are examined in detail, and it provided an in-depth analysis and interpretation of the results in relation to the research questions. The discussion explores how the outcomes of this study align with or diverge from existing literature by shedding light on the broader implications of the research. This chapter also synthesizes the key conclusions drawn from the findings by emphasizing their significance and contributions to the field. Furthermore, practical implications and applications of the research are discussed, and insights into how these findings can be utilized in real-world contexts are offered. Finally, the chapter presents suggestions for future research, and identifies potential areas for further investigation that could build on the current study's findings. This comprehensive discussion aims to provide a clear and cohesive understanding of the study's impact and its relevance to both academic and practical domains.

5.1. Discussion of the Findings

5.1.1. Discussion of Research Question 1

Research Question 1.1.

Is there a significant difference between the gain scores of feedback provided by teachers, peers, and ChatGPT on students' essay scores?

The Research Question 1.1 explored whether there is a significant difference between the gain scores of feedback provided by teachers, peers, and ChatGPT on students' essay scores. As a result, significant differences were observed between the gain scores of feedback provided by different sources. Comparison happened in three directions including the Assignment 1 (teacher feedback) versus Assignment 2 (peer feedback), Assignment 2 (peer feedback) versus Assignment 3 (ChatGPT feedback), Assignment 1 (teacher feedback) versus Assignment 3 (ChatGPT feedback).

Regarding teacher feedback, gain scores were found to be the highest among other gain scores, which is coherent with the existing literature. For example, Ruegg (2015) stated that teacher feedback is the most effective in developing students' writing skills. As teachers are the experts in the subject, they are able to provide more comprehensive and detailed feedback (Lee, 2008). Thus, this helps students to understand their lacking points and revise their writing assignments more effectively. Additionally, Hyland (2003) emphasizes that teacher feedback increases students' motivation, leading to more effective learning. These findings underscore the superior impact of teacher feedback in enhancing student writing skills.

On the other hand, comparing teacher feedback with peer feedback, it was found that peer feedback also achieved good increases in the gain scores but when compared to teacher feedback, it was found to be lower. On this issue, Lundstrom and Baker (2009) state that although peer feedback enables students to think and evaluate critically, it might not be detailed and correct enough to cause significant improvements. However, teacher feedback, provided with the help of the rubric and its explanations, included a few markings, highlights on the paper while peer feedback included them. *Thus, the question of why students improved more after teacher feedback which was less detailed has aroused.*

When comparing the gain scores after ChatGPT feedback and teacher feedback, there was a significant difference, with the gain scores from teacher feedback being higher than those from ChatGPT feedback. Contrary to some propositions in the literature, such as Banihashem et al. (2024) stating that AI feedback could be as effective and qualified as teacher feedback when used in cooperation with teacher feedback, and Escalante et al. (2023) showing that GenAI feedback could be effective and comparable to human feedback, the current study found that ChatGPT feedback made the least increase in the gain scores. Similarly, while ChatGPT provided pages of detailed feedback, teacher feedback was brief and not as detailed as ChatGPT feedback. Interestingly, the quantity of feedback did not seem to matter. When comparing the gain scores of ChatGPT feedback and peer feedback, the results indicated no statistically significant difference between the two. Both feedback sources contributed to improvements in students' essay scores, but their effectiveness was

relatively similar. However, comparing ChatGPT feedback and peer feedback, it is again possible to say that ChatGPT feedback was far more detailed than the ChatGPT feedback. *Thus, the question of why participants improved similarly after both of them has aroused.*

Overall, the findings indicate that while all three types of feedback contributed to student performance improvements, teacher feedback was the most effective, followed by peer feedback, and then GenAI feedback. Comparing the quantity of the feedback from the feedback providers, the ordering from the most detailed to the least detailed is like ChatGPT, peers and teacher. However, when the gain scores were ordered from the most increased to the least increased is the vice versa starting from the teacher, peer and ChatGPT. To find an answer to this issue, the quality of the feedback is needed to be investigated.

Research Question 1.2.

Is there a significant difference between the pre-scores given by the teacher and the scores given by peers and GenAI?

The Research Question 1.2 explored whether there is a significant difference between the pre-scores given by the teacher and the scores given by peers and ChatGPT on students' essays. The analysis revealed notable differences in the initial scores provided by these different feedback sources. As the teacher pre-score was found to be reliable with interrater reliability, the comparison was conducted in two directions. The reason behind doing this comparison is to answer the question aroused on the Research Question 1.1.: although teacher feedback was the less detailed compared to the peer and ChatGPT feedback, the most increase was observed on it. The aim on this research question was to find out whether the feedback provided by the peers and ChatGPT are less qualified compared to the teacher scorings. The teacher pre-scorings, found to be reliable, were compared with the pre-scorings of the other feedback providers.

On the first comparison, the results indicated a statistically significant difference between the mean value of the peer scores for Assignment 2 and the pre-scores of the teacher for Assignment 2, with peer scores being higher than the teacher pre-scores. This suggests that while peer feedback can be valuable for certain aspects of the learning process, it may not always align with the rigorous standards applied by teachers. This finding aligns with the results of Montgomery and Baker (2007), who found that teacher feedback is often more comprehensive and reliable compared to peer feedback. Additionally, Rollinson (2005) highlights that peer feedback may suffer from inconsistencies due to the varying skill levels of peers. From this perspective, it can be stated that peer feedback, being more detailed than teacher feedback, was also lacking in terms of quality compared to teacher feedback. Thus, the lesser gain score increase can be explained here because between the teacher pre-scoring and the peer scoring, there is a significant difference.

On the second comparison, the results indicated no statistically significant difference between the GenAI score for Assignment 3 and the pre-scores of the teacher for Assignment 3. From these results, it can be said that ChatGPT's score, and the teacher's score are similar to each other. However, although the scores are close to each other, a greater increase in the gain scores after teacher feedback was observed. This raises the question of why there is a significant difference between the gain scores after teacher feedback and ChatGPT feedback. To shed light on this issue, qualitative data are investigated in Research Question 3.

In summary, the findings from Research Question 1 highlight significant differences in the effectiveness of feedback provided by teachers, peers, and ChatGPT on students' essay scores. Teacher feedback was found to be the most effective, leading to the highest gain scores and demonstrating its superiority in improving students' writing skills.

5.1.2. Discussion of the Research Question 2

How does the feedback provided by the teacher, peers, and artificial intelligence affect the improvement of students' assignment grades, as measured by pre- and post-feedback scores?

The second research question examined the impact of feedback from teachers, peers, and artificial intelligence on the improvement of students' assignment grades, specifically focusing on pre- and post-feedback scores. The findings from the analysis highlighted the differential effectiveness of each feedback source, emphasizing how various elements within the rubric were influenced by the type of feedback provided.

On the Assignment 1, where participants received feedback from teacher, total mean score for 20 rubric elements in total increased from 3.27 to 4.03, indicating a high improvement. Notable improvements were observed in elements such as complexity and formality level as also suggested in the literature stating the teachers' feedback positive effect on the structure of the students' written works. On the other hand, elements such as thesis statement and citations remained unchanged. This significant improvement in Assignment 1 suggests that teacher feedback is highly effective in enhancing students' performance. From the literature, Razi (2016) emphasizes the effectiveness of multiple feedback sources, including teacher feedback, in enhancing academic writing quality and reliability. This supports the finding that teacher feedback significantly improves students' performance on assignments by enhancing rubric elements such as complexity and formality. Similarly, Hyland and Hyland (2019) highlight the importance of clear, specific, and timely teacher feedback. They state that well-delivered teacher feedback can build students' confidence and skills, which aligns with the observed improvement in Assignment 1. Thus, teacher feedback proved to be the most effective in enhancing students' assignment grades.

On the Assignment 2, where participants received feedback from the peers, total mean score for 20 rubric elements in total increased from 3.09 to 3.26, indicating a moderate

improvement. Rollinson (2005) discusses the benefits of structured peer feedback in ESL writing classrooms, noting significant improvements in writing quality when peer feedback is well-organized and supported by training. Notable improvements were observed in elements such as thesis statement, paragraph, concluding remarks. Some elements, like citations, showed no improvement. Therefore, peer feedback, while beneficial, was less effective than teacher feedback in improving assignment grades. Existing body of literature supports this finding by Cui, et al. (2022) stating that while teacher feedback led to more revisions, peer feedback fostered greater student autonomy and engagement. This aligns with the finding that peer feedback in Assignment 2 led to moderate improvements compared to teacher feedback. Furthermore, some studies, like Nelson and Schunn (2009), highlight that peer feedback can complement but not fully replace teacher feedback due to variability in feedback quality and students' ability to critique effectively. This could suggest a limitation in the effectiveness of peer feedback as noted in Assignment 2.

On the Assignment 3, where participants received feedback from ChatGPT, total mean score for 20 rubric elements in total increased from 3.59 to 3.75, indicating a moderate improvement. Notable improvements were observed in elements such as paragraph development, grammar and punctuation and formality level. Supporting this finding, Godwin-Jones (2022) reviewed multiple studies and noted that AI tools are effective for grammatical and lexical errors. However, some elements, like thesis statement and citations, showed minimal or no improvement as also stated by Godwin-Jones stating that ChatGPT feedback falls short in addressing higher-order writing skills such as organization and argumentation. This suggests potential limitations in the effectiveness of AI feedback observed in the study. Hence, AI feedback can be beneficial, but this benefit could vary. This variation was explained by Taiye et al. (2024) examining the impact of Generative AI tools and stating that while AI tools provided valuable feedback, their effectiveness varied based on the interaction level and student proficiency, which aligns with the mixed results observed in Assignment 3. Similarly, Nguyen et al. (2024) found that interactive use of AI tools like ChatGPT enhances performance compared to using them as supplementary sources, supporting the findings that AI feedback can be beneficial for specific elements like paragraph development and grammar. Thus, ChatGPT feedback could be used as a supplementary to other feedback.

The literature review findings generally align with existing studies on the effectiveness of teacher, peer, and AI feedback. Teacher feedback is consistently shown to be the most effective, peer feedback provides moderate benefits, and AI feedback, while beneficial, has limitations in certain areas.

5.1.3. Discussion of the Research Question 3

How do students perceive the strengths and weaknesses of feedback provided by teacher, peers, and GenAI?

In this section, the perceptions of students regarding the strengths and weaknesses of feedback provided by teachers, peers, and GenAI (ChatGPT) will be examined. Research Question 3 aims to understand how these different types of feedback are perceived by students and their impact on the student experience. Student perceptions of feedback can significantly influence the effectiveness of the feedback and the subsequent improvements in their writing skills. Also, in previous sections, several unresolved issues emerged that quantitative data alone could not fully explain. For instance, it was observed that teacher feedback, although less detailed, resulted in higher gain scores compared to more detailed peer and ChatGPT feedback. Additionally, even though there were no significant differences between the pre-scores given by the teacher and ChatGPT, ChatGPT feedback did not lead to proportional improvements in student performance. These discrepancies indicate that qualitative aspects of feedback perception play a crucial role in understanding these outcomes.

Teacher feedback was valued for its trustworthiness and brevity. Students appreciated the clear, concise feedback which allowed them to understand their strengths and weaknesses quickly. This was echoed in the literature stating that teacher feedback is valued by the students if the feedback includes outlines of their strengths and weaknesses (Moser, 2020) However, the primary weakness was the lack of detail, often attributed to the teacher's limited time, which was also echoed in the literature as the high demands on the

teachers by Lee (2017) leading to some mistakes being overlooked and resulting in superficial corrections.

Peer feedback was praised for its detail and motivational aspects. Detailed comments provided diverse perspectives, and the supportive environment reduced fear of judgment. However, it was often criticized for being incorrect, inconsistent, and untrustworthy. In the literature, similar drawbacks of the peer feedback were also mentioned including a lack of feedback skills among peers (Li et al., 2012), a task-oriented mindset to please the teacher (Kaufmann and Schunn, 2011), lack of trust among students (Lu and Law, 2012), avoidance of negative comments (Liu et al., 2020), inhibition in L2 expression (Lundstrom and Baker, 2009), and superficial feedback (Patchan et al., 2009).

GenAI feedback was noted for its detailed and comprehensive nature, high availability, and extensive capability. This was also echoed in the literature with the studies by Godwin-Jones (2022) stating that AI-based feedback tools are effective in providing accurate and timely feedback on grammatical and lexical errors and by Escalante et al. (2023) stating that the participants appreciated the clarity and specificity of AI-generated feedback. However, it faced criticism for occasional unreliability, mechanical interaction, and deceptive scoring. Students reported instances of incorrect information and felt the lack of a human touch was a drawback. These issues were also mentioned in the literature by Gayed et al. (2022) noting that the AI feedback needs careful consideration and human oversight.

In conclusion, this study underscores the varying effectiveness, and perceptions of feedback provided by teachers, peers, and GenAI in enhancing students' essay scores. Teacher feedback emerged as the most effective, offering concise yet impactful guidance, despite its brevity. This aligns with existing literature that highlights the expertise and motivational impact of teacher feedback, making it a reliable source for improving student performance. However, the brevity of teacher feedback raises concerns about its potential to overlook detailed corrections. Peer feedback, while detailed and motivational, was found to be less reliable and consistent. The detailed nature of peer feedback provided diverse perspectives and fostered a supportive environment, yet the inconsistencies and inaccuracies

posed significant challenges. This echoes findings in the literature that peer feedback can suffer from variability due to differing skill levels among peers. GenAI feedback, characterized by its comprehensiveness and high availability, showed limitations in reliability, and lacked a personal touch. The detailed and timely feedback provided by AI tools was appreciated by students, but the occasional unreliability and mechanical nature of interactions posed significant drawbacks. This suggests that while AI can supplement human feedback, it cannot fully replace the nuanced and empathetic feedback provided by teachers. Overall, these findings indicate that while all three types of feedback contribute to student performance improvements, teacher feedback remains paramount due to its reliability and motivational impact.

5.2. Conclusions

This study aimed to investigate which type of feedback including peer, teacher and GenAI most significantly improved the students' writing performance. Additionally, the study explored the consistency and reliability of the scores assigned by different evaluators by focusing on possible discrepancies or agreements among the scores given by the teacher, peers and GenAI. By exploring this, the study shed light on whether it is possible to use GenAI for assessment or not. Lastly, the study unveiled the students' perceptions of the feedback they received from the teacher, peers and GenAI on their weaknesses and strengths.

To continue by these objectives, the study adopted the mixed method approach with instrumental collective case study design. This study was conducted with 16 students out of 103 students who successfully completed Writing Skills I course during the fall semester of the first year at the Department of English Language Teaching, Faculty of Education, at a state university in Türkiye. For the Research Question 1.1., data sources included the gain scores of the Assignment 1, 2, and 3 collected with quantitative methodology by FAWR and analysed with Non-parametric Wilcoxon Signed Ranks Test. For the Research Question 1.2., data sources included the pre-scores of the Assignment 2 and 3 from the teacher and the feedback scores from peers and GenAI and 3 collected with quantitative methodology by FAWR and analysed with Non-parametric Wilcoxon Signed Ranks Test. For the Research

Question 2, data sources included the gain scores of the Assignment 1, 2, and 3 collected with quantitative methodology by FAWR and analysed with Non-parametric Wilcoxon Signed Ranks Test. For the Research Question 2., data sources included the pre- and post-scores of Assignment 1,2 and 3 collected with quantitative and qualitative methodology by FAWR and analysed with narration and descriptive statistics. For the Research Question 2, data sources included the gain scores of the Assignment 1, 2, and 3 collected with quantitative methodology by FAWR and analysed with Non-parametric Wilcoxon Signed Ranks Test. For the Research Question 3, data sources included the participants from which the data were collected with qualitative methodology by interview questions and analysed with narration and qualitative content analysis. In the light of these processes, analyses and with these aims at hand, the conclusions of this research were presented in items below based on the research questions.

5.2.1. Conclusions from the Research Question 1

- Comparing the gain scores of each assignment after feedback, the gain scores of Assignment 1, which was provided with teacher feedback, were significantly higher than those of Assignments 2 and 3. Specifically, the Wilcoxon Signed Ranks Test results showed significant differences in the gain scores when comparing the scores after teacher feedback) with both scores after peer feedback and scores after GenAI feedback, with scores after teacher feedback consistently showing higher gains. ***This indicates that participants improved themselves from teacher feedback compared to peer and GenAI feedback.***
- When comparing the gain scores of Assignment 2 with Assignment 3, no statistically significant difference was found. ***This suggests that the effectiveness of peer feedback and GenAI feedback was relatively similar in terms of participants' progress. This also indicates that either method can be a viable support when teacher feedback is not available.***
- There was a significant difference between the pre-scores given by the teacher and the scores given by peers for the Assignment 2 showing peer scores being higher. ***This could imply that peers tended to be more tolerant or optimistic in their scoring***

compared to the teacher. This might reflect a more generous or less critical perspective.

- No significant difference was found between the pre-scores given by the teacher and the scores given by GenAI for Assignment 3. ***This indicates that the scoring provided by GenAI was relatively aligned with the teacher's pre-scores, suggesting a level of consistency between the teacher and GenAI evaluations. This shows that GenAI scoring aligns more closely with teacher pre-scores, suggesting that GenAI can provide an objective and consistent evaluation.***

From the findings and discussion of Research Question 1, feedback from human teachers is indispensable and irreplaceable for students' academic writing skills. However, the objective and consistent feedback provided by GenAI can support and complement educational processes. Combining human guidance and intuition with the analytical capabilities of AI can lead to more comprehensive and effective educational outcomes. Given that students showed similar improvements after receiving feedback from peers and GenAI, it can be inferred that peer and GenAI feedback may serve as effective substitutes. However, peer feedback was less consistent with teacher scores compared to GenAI feedback. This suggests that while peer feedback is valuable, GenAI provides a level of consistency and alignment with teacher evaluations, which is beneficial for maintaining standardization and objectivity in assessments.

5.2.2. Conclusions from the Research Question 2

- Teacher feedback resulted in substantial improvements in multiple rubric elements across assignments including the elements such as the thesis statement, citations, and grammar showing notable increases in scores. ***This indicates that teacher feedback was effective in addressing specific areas of weakness and guiding students towards better performance.***

- The greatest improvements were observed in areas such as citation needed, complexity, grammar, and concluding remarks. ***This suggests that teacher feedback was particularly impactful in enhancing these aspects of writing.***
- Peer feedback led to moderate improvements in some rubric elements, particularly in the second assignment on the elements such as citations, direct quotations, and paragraph development seeing slight increases. ***This indicates that peer feedback was somewhat helpful in improving these areas.***
- The effectiveness of peer feedback varied across participants and assignments, with some participants showing significant gains while others did not exhibit any improvement. ***This variability highlights the inconsistency in the quality and usefulness of peer feedback.***
- GenAI feedback resulted in improvements in several rubric elements, particularly in the third assignment including the elements such as formality level, grammar, and punctuation seeing increases in scores. ***This indicates that GenAI feedback was effective in enhancing technical aspects of writing.***
- While GenAI feedback showed consistency in improving certain elements, it was less effective in others, such as paragraph development and the flow of ideas. ***This suggests that GenAI feedback may be more suited to addressing specific, well-defined aspects of writing rather than more complex, holistic elements.***

From the findings and discussion of Research Question 2, teacher feedback consistently led to the most significant improvements across assignments, indicating its critical role in guiding students towards better performance while peer feedback, sometimes helpful, showed variability in effectiveness, suggesting that its impact depends on the quality and engagement of the peers providing and receiving the feedback. On the other hand, GenAI feedback proved to be a useful tool for improving technical aspects of writing but was less effective in addressing more complex, integrative elements of assignments. In conclusion, the combination of feedback from teachers, peers, and GenAI can provide a comprehensive support system for students. Teacher feedback can address major weaknesses, peer feedback can offer diverse perspectives, and GenAI feedback can enhance technical accuracy.

5.2.3. Conclusions from the Research Question 3

- Teacher feedback is highly valued for its trustworthiness thanks to the teacher's expertise in academic writing and familiarity with course requirements. Students find this feedback reliable and accurate, ***which significantly helps their learning and performance improvement.***
- The concise nature of the feedback from the teacher allows students to clearly understand the strengths and weaknesses of their work without ambiguity, ***which enabled them to take direct and well-guided actions. This brevity also helps in understanding the assessment system, making it easier for students to identify areas needing improvement.***
- Although less frequent, teacher feedback can be motivating. ***It encourages students to work for higher scores and correct their mistakes and it also demonstrated the teacher's commitment to their improvement.***
- The most significant issue with teacher feedback is its lack of detail, often due to time constraints of the teacher. ***This can result in superficial feedback that may overlook some mistakes.***
- Some students find the feedback difficult to understand, especially because they received it for the first time. ***This can lead to confusion and hinder the effective use of the feedback.***
- Feedback can sometimes be disheartening, particularly when students believe they performed well and expect to receive high scores or comments. ***This emotional response can affect their perception of the feedback's value.***
- Peer feedback is often detailed due to course requirements, ***which motivated students to provide comprehensive and thorough evaluations.***
- Receiving feedback from peers can be less intimidating and more encouraging, fostering a supportive learning environment. ***It can boost students' confidence and encourage them to improve their work.***

- Peer feedback can be prone to errors and inconsistencies, as peers are still learning and may provide conflicting or misguided advice. ***This can lead to confusion and frustration among the peers receiving the feedback.***
- The variability in the quality of peer feedback makes it less reliable. Some peers may give feedback negligently with the aim of fulfilling the requirement without genuine effort, ***which diminishes its effectiveness.***
- GenAI provides highly detailed feedback by identifying even minor errors and offering comprehensive suggestions. ***This level of detail is unmatched by other feedback sources.***
- GenAI is always available, and it allows students to receive feedback whenever needed. It can handle multiple requests simultaneously, ***acting as a private tutor.***
- When the feedback is accurate, GenAI offers reliable insights supported by specific examples from the students' work.
- Despite its strengths, GenAI can provide incorrect or misleading information due to its ongoing development. ***This unreliability can undermine students' trust in the feedback.***
- The lack of a human touch in GenAI feedback makes interactions feel artificial. Students miss the personalized and empathetic engagement they get from human feedback providers.
- GenAI can sometimes provide overly generous scores for certain elements, which may mislead students about their actual performance and necessary improvements.

These insights suggest a complementary approach to feedback provision. Combining the trustworthiness and clarity of teacher feedback, the detail and motivational aspects of peer feedback, and the comprehensive and available nature of GenAI feedback could provide students with a well-rounded and effective feedback experience. Addressing the weaknesses of each feedback type, such as improving the detail in teacher feedback and ensuring the reliability of GenAI feedback, would further enhance the overall feedback process for students.

5.2.4. Ethical Use of Generative Artificial Intelligence in the Process of Academic Writing

This study also highlighted the importance of the ethical use of GenAI tools. By integrating developed and tested prompts formulated within the boundaries of ethical usage, educators can guide students to use GenAI responsibly and effectively. These prompts ensure that students understand the boundaries of ethical use and the significance of maintaining academic honesty in their work. The integration of GenAI in academic writing has significant implications for maintaining academic integrity. This study underscores the necessity of ethical guidelines and structured prompts to guide the responsible use of GenAI tools. By incorporating these elements, educators can help students understand the importance of using AI responsibly and maintaining academic honesty. The conclusions from this research indicate that the ethical use of GenAI can enhance learning experiences while upholding the standards of academic integrity.

5.3. Implications

In the light of the conclusions drawn from this study, several important implications are offered for educational practices, policy practices, implications for students and implications for academic integrity.

5.3.1. Educational Implications

- Given the critical role of teacher feedback in significantly improving student performance, it is essential to reduce the workload on teachers to allow them to provide detailed and frequent feedback. Strategies could include reducing the number of students per class, hiring additional teaching assistants, and providing administrative support to free up teachers' time for feedback.
- Although peer feedback shows variability in effectiveness, it remains valuable for its detailed nature and supportive environment. Incorporating structured peer feedback

sessions and training students in effective feedback techniques can enhance its reliability and impact.

- The use of GenAI for feedback can be a valuable supplement, especially for technical aspects of writing. Integrating GenAI tools into the curriculum can provide students with immediate and detailed feedback, complementing the feedback from teachers and peers.
- The study suggests that a combination of teacher, peer, and GenAI feedback provides a comprehensive support system for students. Educational institutions should consider adopting hybrid feedback models that leverage the strengths of each feedback type to offer a more balanced and effective learning experience.

5.3.2. Policy Implications

- Regarding the resources, policymakers should allocate resources to support the integration of GenAI tools in educational settings, ensuring that both students and teachers have access to these technologies. This includes funding for software, training, and ongoing support.
- The alignment of GenAI feedback with teacher evaluations suggests that GenAI can play a role in standardizing assessments. Policymakers could explore setting guidelines for the use of GenAI in grading and feedback processes to ensure consistency and objectivity across educational institutions.
- Students can benefit from understanding the strengths and limitations of different feedback types. Educators should guide students on how to effectively use and integrate feedback from teachers, peers, and GenAI to improve their academic writing skills comprehensively.
- Developing students' feedback literacy is crucial. Training programs and workshops can help students become better at giving and receiving feedback, making the process more productive and less intimidating.

5.3.3. Student Implications

- Students can benefit from understanding the strengths and limitations of different feedback types. Educators should guide students on how to effectively use and integrate feedback from teachers, peers, and GenAI to improve their academic writing skills comprehensively.
- Developing students' feedback literacy is crucial. Training programs and workshops can help students become better at giving and receiving feedback, making the process more productive and less intimidating.
- Establishing clear guidelines and training on the ethical use of GenAI can provide students with a framework to understand the importance of academic honesty and the potential pitfalls of misuse. This can be incorporated into the curriculum as part of digital literacy education.
- Organizing workshops and training sessions focused on the ethical use of GenAI and academic integrity can enhance students' awareness and understanding. These initiatives can teach students how to leverage GenAI tools ethically and responsibly.
- Even when structured prompts are provided, the responsibility of making ethical decisions lies with the students. They must manage how they direct the prompts, interpret the responses, and use the content generated by GenAI. This underscores the need for comprehensive academic integrity education.
- Students must critically evaluate AI-generated content to ensure it does not contain academic misconduct. They should understand that it is their responsibility to verify the accuracy and originality of the information provided by GenAI.
- Providing students with thorough education on academic integrity is crucial. This training should emphasize the importance of ethical decision-making, proper citation practices, and the potential consequences of academic misconduct. By understanding these principles, students can better navigate the challenges of using AI tools in their academic work.

5.3.4. Academic Integrity Implications

- The creation of structured prompts that include ethical usage guidelines can serve as a crucial educational resource. These prompts can help students navigate the ethical considerations of using AI in their academic work.
- Establishing clear guidelines and training on the ethical use of GenAI can provide students with a framework to understand the importance of academic honesty and the potential pitfalls of misuse. This can be incorporated into the curriculum as part of digital literacy education.
- Embedding ethical considerations into assessment criteria can encourage students to practice academic integrity consistently. This can include evaluating students on their adherence to ethical guidelines when using GenAI tools.

5.4. Suggestions for Further Research

This study's findings suggest several future research directions and AI development and integration recommendations.

- As the peer feedback was found inaccurate and inconsistent by the participants, future research studies could focus on exploring ways to improve the accuracy and consistency of peer feedback building on this study.
- Not for research studies but for AI tool developers, developing and testing advanced AI feedback systems that can better mimic the nuanced feedback provided by human experts is another suggestion aroused from the findings of this study.
- Furthermore, as this study inspected a period of time that participants integrated AI, investigating the long-term effects of AI-generated feedback on student writing skills could be another research study to explore the more prolonged effects of AI on writing skills. This could also include how continuous use of AI feedback impacts the development of critical thinking and writing proficiency over multiple semesters or academic years.

- Additionally, as this study was conducted in a specific department, conducting comparative studies to evaluate the effectiveness of AI feedback in different academic disciplines could show the different effects of AI.
- Explore the development and implementation of more advanced AI algorithms that can provide not only grammatical and lexical corrections but also high-level structural and content-related feedback. Research could focus on how these enhancements affect the overall quality and usefulness of AI feedback.
- Investigating the potential for collaborative feedback mechanisms where AI and human teachers work together to provide a cohesive and comprehensive feedback experience is another future research suggestion aroused from this study. For example, AI could provide preliminary feedback that teachers then refine and expand upon. Research could assess the efficiency and effectiveness of such hybrid models and also develop frameworks for combining these feedback sources in ways that maximize student learning and engagement.
- Study how students' perceptions of AI feedback evolve over time and what factors influence their acceptance and utilization of AI tools. This could include examining how different demographic factors such as age, tech-savviness impact students' benefitting from AI feedback.
- Conduct research on how AI feedback affects students from diverse backgrounds, including non-native English speakers, students with different learning styles, and those with varying levels of prior academic achievement like students with degrees of Master of Arts, or Doctor of Philosophy, or with preparation classes.

5.5. The Summary of the Chapter

In Chapter Five, the findings from the study are meticulously analysed and discussed in relation to the research questions, drawing connections with existing literature. The discussion reveals significant insights into the effectiveness of feedback from teachers, peers, and GenAI (ChatGPT) on students' essay scores. Teacher feedback was found to be the most effective, enhancing students' performance significantly more than peer or GenAI

feedback. Peer feedback, while valuable for its detail and motivational support, showed variability and inconsistencies in its effectiveness. GenAI feedback, though comprehensive and highly detailed, was less effective in improving higher-order writing skills but provided consistent technical corrections.

Students' perceptions highlighted the strengths and weaknesses of each feedback type. Teacher feedback was valued for its trustworthiness and motivational impact but was often limited by its brevity. Peer feedback was praised for its detail but criticized for inconsistency and unreliability. GenAI feedback was appreciated for its availability and detail but faced challenges due to occasional inaccuracies and lack of personal touch.

The study concludes that while teacher feedback remains paramount, integrating peer and GenAI feedback can provide a more balanced and effective feedback system. Practical implications include reducing teacher workload, enhancing peer feedback training, and integrating GenAI tools into curricula. Ethical considerations are emphasized, particularly in the use of GenAI, with suggestions for structured prompts and guidelines to ensure academic integrity. The chapter suggests future research directions, such as improving peer feedback accuracy, developing advanced AI algorithms, and exploring long-term effects of AI feedback on student writing skills.

REFERENCES

- Alharbi, W. (2023). "AI in the foreign language classroom: A pedagogical overview of automated writing assistance tools". *Education Research International*, 4253331. <https://doi.org/10.1155/2023/4253331>
- Badger, R., and White, G. (2000). "A process genre approach to teaching writing". *ELT Journal*, 54(2), 153-160. <https://doi.org/10.1093/elt/54.2.153>
- Bailey, S. (2003). *Academic writing: A handbook for international students*. Routledge. <https://doi.org/10.4324/9780203464120>
- Banihashem, S. K., Kerman, N. T., and Noroozi, O. (2024). "Feedback sources in essay writing: Peer-generated or AI-generated feedback?". *International Journal of Educational Technology in Higher Education*, 21, 23. <https://doi.org/10.1186/s41239-024-00455-4>
- Barrett, A., and Pack, A. (2023). "Not quite eye to A.I.: Student and teacher perspectives on the use of generative artificial intelligence in the writing process". *International Journal of Educational Technology in Higher Education*, 20, 59. <https://doi.org/10.1186/s41239-023-00427-0>
- Beauvois, M. H. (1998). "Conversations in slow motion: Computer-mediated communication in the foreign language classroom". *Canadian Modern Language Review*, 54, 198-217. <https://doi.org/10.3138/cmlr.54.2.198>
- Bengio, Y., Hinton, G., and LeCun, Y. (2019). "Deep learning". *Nature*, 521, 436-444. <https://doi.org/10.1038/nature14539>
- Bharadwaj, P., Shaw, C., NeJame, L., Martin, S., Janson, N., and Fox, K. (2023, June). *Time for class - 2023*. Tyton Partners. Retrieved on 10 August 2023 from https://tytonpartners.com/app/uploads/2023/06/Time-for-Class-2023-Report_Final.pdf
- Bitchener, J. (2008). "Evidence in support of written corrective feedback". *Journal of Second Language Writing*, 17, 102-118. <http://dx.doi.org/10.1016/j.jslw.2007.11.004>
- Bitchener, J., and Knoch, U. (2008). "The value of a focused approach to written corrective feedback". *ELT Journal*, 63(3), 204-211. <https://doi.org/10.1093/elt/ccn043>

- Bitchener, J., and Knoch, U. (2010). "The contribution of written corrective feedback to language development: A ten month investigation". *Applied Linguistics*, 31, 193-214. <https://doi.org/10.1093/applin/amp016>
- Bitchener, J., Young, S., and Cameron, D. (2005). "The effect of different types of corrective feedback on ESL student writing". *Journal of Second Language Writing*, 14, 191-205. <https://doi.org/10.1016/j.jslw.2005.08.001>
- Blake, R. J., and Zyzik, E. C. (2003). "Who's helping whom? Learner/heritage-speakers' networked discussions in Spanish". *Applied Linguistics*, 24(4), 519-544. <https://doi.org/10.1093/applin/24.4.519>
- Bozkurt, A. (2023a). "Generative artificial intelligence (AI) powered conversational educational agents: The inevitable paradigm shift". *Asian Journal of Distance Education*, 18(1), 198-204. <https://doi.org/10.5281/zenodo.7716416>
- Bozkurt, A. (2023b). "Unleashing the potential of generative AI, conversational agents and chatbots in educational praxis: A systematic review and bibliometric analysis of GenAI in education". *Open Praxis*, 15(4), 261-270. <https://doi.org/10.55982/openpraxis.15.4.609>
- Bozkurt, A. (2024). "GenAI et al.: Co-creation, authorship, ownership, academic ethics and integrity in a time of generative AI". *Open Praxis*, 16(1), 1-10. <https://doi.org/10.55982/openpraxis.16.1.654>
- Bretag, T., Mahmud, S., East, J., Green, M., James, C., McGowan, U., Partridge, L., Walker, R., and Wallace, M. (2011). Academic integrity standards: A preliminary analysis of the academic integrity policies at Australian universities. Paper presented at the Australian Universities Quality Forum (AUQF), Melbourne, 29 June–1 July 20. Retrieved on 10 August 2023 from <https://ro.uow.edu.au/asdpapers/323>
- Brown, H. D. (2000). *Principles of language learning and teaching* (4th ed.). Pearson Education.
- Bruner, J. S., and Sherwood, V. (1976). Early rule structure: The case of "peekaboo." In R. Harre (Ed.), *Life sentences: Aspects of the social role of language*. Wiley.
- Chan, C. K. Y., and Hu, W. (2023). "Students' voices on generative AI: Perceptions, benefits, and challenges in higher education". *International Journal of Educational Technology in Higher Education*, 20, 43. [https://doi.org/10.1186/s41239-023-00411-](https://doi.org/10.1186/s41239-023-00411-8)

- Chen, B., Zhang, Z., Langren'e, N., and Zhu, S. (2023). "Unleashing the potential of prompt engineering in large language models: A comprehensive review". *Arxiv preprint*. <https://doi.org/10.48550/arXiv.2310.14735>
- Chiu, T. K. F. (2023). "The impact of generative AI (GenAI) on practices, policies and research direction in education: A case of ChatGPT and Midjourney". *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2023.2253861>
- Chiu, T. K. F. (2024). "Future research recommendations for transforming higher education with generative AI". *Computers and Education: Artificial Intelligence*, 6, 100197. <https://doi.org/10.1016/j.caeai.2023.100197>
- Coe, R. M. (2002). The new rhetoric of genre: Writing political briefs. In A. Johns (Ed.), *Genre in the classroom: Multiple perspectives* (pp. 197-207). Erlbaum.
- Committee on Publication Ethics. (2023, February 13). Authorship and AI tools. Retrieved on 10 August 2023 from <https://publicationethics.org/cope-position-statements/ai-author>
- Cotton, D. R. E., Cotton, P. A., and Shipway, J. R. (2023). "Chatting and cheating: Ensuring academic integrity in the era of ChatGPT". *Innovations in Education and Teaching International*, 61(2), 228-239. <https://doi.org/10.1080/14703297.2023.2190148>
- Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Pearson Education.
- Crews, T., and Butterfield, J. B. (2014). "Data for flipped classroom design: Using student feedback to identify the best components from online and face-to-face classes". *Higher Education Studies*, 4, 38-47. <https://doi.org/10.5539/hes.v4n3p38>
- Cribb, M. (2002). Respecting process and product in research papers (Unpublished doctoral dissertation). Kansai Gaidai University.
- Cui, Y., Schunn, C. D., and Gai, X. (2022). "Peer feedback and teacher feedback: A comparative study of revision effectiveness in writing instruction for EFL learners". *Higher Education Research and Development*, 41(6), 1838-1854. <https://doi.org/10.1080/07294360.2021.1969541>
- De Florio, I. (2023). *From Assessment to Feedback: Applications in the Second/Foreign Language Classroom*. Cambridge University Press.
- DiGiovanni, E., and Nagaswami, G. (2001). "Online peer review: An alternative to face-to-face?". *ELT Journal*, 55(3), 263-272. <https://doi.org/10.1093/elt/55.3.263>

- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.
- Eaton, S. E. (2023). Comprehensive academic integrity (CAI): An ethical framework for educational contexts. In S. E. Eaton (Ed.), *Handbook of Academic Integrity*. Springer. https://doi.org/10.1007/978-981-287-079-7_194-1
- Ellis, R. (2009). "Corrective feedback and teacher development". *L2 Journal*, 1, 3-18. <https://doi.org/10.5070/L2.V1I1.9054>
- Ellis, R., Sheen, Y., Murakami, M., and Takashima, H. (2008). "The effects of focused and unfocused written corrective feedback in an English as a foreign language context". *System*, 36, 353-371. <https://doi.org/10.1016/j.system.2008.02.001>
- Ene, E., and Upton, T. A. (2018). "Synchronous and asynchronous teacher electronic feedback and learner uptake in ESL composition". *Journal of Second Language Writing*, 41, 1-13. <https://doi.org/10.1016/j.jslw.2018.05.005>
- Escalante, J., Pack, A., and Barrett, A. (2023). "AI-generated feedback on writing: Insights into efficacy and ENL student preference". *International Journal of Educational Technology in Higher Education*, 20, 57. <https://doi.org/10.1186/s41239-023-00425-2>
- Escotet, M. Á. (2023). "The optimistic future of artificial intelligence in higher education". *Prospects*. <https://doi.org/10.1007/s11125-023-09642-z>
- European Commission, Directorate-General for Education, Youth, Sport and Culture. (2022). Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators. Publications Office of the European Union. Retrieved on 10 August 2023 from <https://data.europa.eu/doi/10.2766/153756>
- Feigenbaum, E. A., and Reddy, R. (1994). "Artificial intelligence: What have we learned, and where are we going?". *Artificial Intelligence*, 59(1-2), 1-30.
- Ferris, D. (2002). *Treatment of error in second language student writing*. University of Michigan Press.
- Ferris, D. R. (2003). *Response to student writing: Implications for second language students*. Lawrence Erlbaum Associates. <https://doi.org/10.4324/9781410607201>
- Ferris, D. R. (2014). "Responding to student writing: Teachers' philosophies and practices". *Assessing Writing*, 19, 6-23. <https://doi.org/10.1016/j.asw.2013.09.004>

- Ferris, D., and Roberts, B. (2001). "Error feedback in L2 writing classes: How explicit does it need to be?". *Journal of Second Language Writing*, 10, 161-184. [https://doi.org/10.1016/S1060-3743\(01\)00039-X](https://doi.org/10.1016/S1060-3743(01)00039-X)
- Flavell, J. H. (1979). "Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry". *American Psychologist*, 34, 906-911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Fong, C. J., Williams, K. M., Williamson, Z. H., Lin, S., Kim, Y. W., and Schallert, D. L. (2018). "'Inside out': Appraisals for achievement emotions from constructive, positive, and negative feedback on writing". *Motivation and Emotion*, 42(2), 236-257. <https://doi.org/10.1007/s11031-017-9658-y>
- Freedman, A., and Medway, P. (1994). *Genre and the new rhetoric*. Guilford Press.
- Gabrielatos, C. (2002). "EFL writing: Product and process". (ED476839). Retrieved on 10 August 2023 from <http://files.eric.ed.gov/fulltext/ED476839.pdf>
- Gallant, T. B. (2008). Academic integrity in the twenty-first century: A teaching and learning imperative. *ASHE Higher Education Report*, 33(5). <http://dx.doi.org/10.1002/aehe.3305>
- Gayed, J. M., Carlon, M. K. J., Oriola, A. M., and Cross, J. S. (2022). "Exploring an AI-based writing assistant's impact on English language learners". *Computers and Education: Artificial Intelligence*, 3, 100055. <https://doi.org/10.1016/j.caeai.2022.100055>
- Gielen, S., Tops, L., Dochy, F., Onghena, P., and Smeets, S. (2010). "A comparative study of peer and teacher feedback and of various peer feedback forms in a secondary school writing curriculum". *British Educational Research Journal*, 36(1), 143–162. Retrieved on 10 August 2023 from <http://www.jstor.org/stable/27823591>
- Giray, L. (2023). "Prompt engineering with ChatGPT: A guide for academic writers". *Annals of Biomedical Engineering*, 51, 2629–2633. <https://doi.org/10.1007/s10439-023-03272-4>
- Godwin-Jones, R. (2022). "Partnering with AI: Intelligent writing assistance and instructed language learning". *Language Learning and Technology*, 26(2), 5–24. <http://doi.org/10125/73474>

- Goldstein, L. M. (2004). "Questions and answers about teacher written commentary and student revision: Teachers and students working together". *Journal of Second Language Writing*, 13(1), 63-80. <https://doi.org/10.1016/j.jslw.2004.04.006>
- Goldstein, L., and Hyland, K. (2006). Feedback and revision in second language writing: Contextual, teacher, and student variables. In K. Hyland and F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 185-205). Cambridge University Press.
- Gönülal, T., and Loewen, S. (2018). Scaffolding technique. In J. I. Lontas (Ed.), *The TESOL Encyclopaedia of English Language Teaching*. John Wiley and Sons. <https://doi.org/10.1002/9781118784235.eelt0032>
- Grabe, W., and Kaplan, R. B. (1996). *Theory and practice of writing: An applied linguistic perspective*. Longman.
- Graham, S., and Harris, K. R. (1997). "Self-regulation and writing: Where do we go from here?". *Contemporary Educational Psychology*, 22(1), 102-114. <https://doi.org/10.1006/ceps.1997.0920>
- Gu, J., Han, Z., Chen, S., Beirami, A., He, B., and Zhang, G. (2023). "A systematic survey of prompt engineering on vision-language foundation models". *ArXiv preprint*. <https://doi.org/10.48550/arXiv.2307.12980>
- Guo, Y. (2005). "A process-genre model for teaching writing". *English Teaching Forum*, 43(3), 18-26. Retrieved on 10 August 2023 from https://americanenglish.state.gov/files/ae/resource_files/05-43-3-d.pdf
- Hamer, J., Purchase, H., Luxton-Reilly, A., and Denny, P. (2014). "A comparison of peer and tutor feedback". *Assessment and Evaluation in Higher Education*, 40(1), 151–164. <https://doi.org/10.1080/02602938.2014.893418>
- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.
- Ho, M., and Savignon, S. J. (2007). "Face-to-face and computer-mediated peer review in EFL writing". *CALICO Journal*, 24(2), 269-290. Retrieved on 10 August 2023 from <http://www.jstor.org/stable/24147912>
- Hu, G. (2019). Culture and peer feedback. In K. Hyland and F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 45-63). Cambridge University Press.

- Hubbard, P. (2010). An invitation to CALL foundations of computer-assisted language learning. In J. I. Lontas (Ed.), *The TESOL Encyclopaedia of English Language Teaching*. John Wiley and Sons.
- Hyland, K. (2002). *Teaching and researching writing*. Pearson Education.
- Hyland, K. (2003). "Genre-based pedagogies: A social response to process". *Journal of Second Language Writing*, 12, 17-29. [https://doi.org/10.1016/S1060-3743\(02\)00124-8](https://doi.org/10.1016/S1060-3743(02)00124-8)
- Hyland, K. (2003). *Second language writing*. Cambridge: Cambridge University Press.
- Hyland, K., and Hyland, F. (2006). "Feedback on second language students' writing". *Language Teaching*, 39(2), 83-101. <https://doi.org/10.1017/S0261444806003399>
- Hyland, K., and Hyland, F. (2019). Contexts and issues in feedback on L2 writing. In K. Hyland and F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 1-22). Cambridge University Press.
- Hyland, K., and Shaw, P. (2016). Introduction. In K. Hyland and P. Shaw (Eds.), *The Routledge handbook of EAP* (pp. 1-13). Routledge.
- Hyon, S. (1996). "Genre in three traditions: Implications for ESL". *TESOL Quarterly*, 30(4), 693-722. <https://doi.org/10.2307/3587930>
- IBM. (n.d.). What are large language models (LLMs)? IBM. Retrieved on 10 August 2023 from <https://www.ibm.com/topics/large-language-models>
- International Center for Academic Integrity. (n.d.). Statement on academic integrity and artificial intelligence. Retrieved on 10 August 2023 from https://academicintegrity.org/images/ICAI_Statement_on_Academic_Integrity_Artificial_Intelligence.pdf
- Irwin, B. (2017). "Written corrective feedback: Student preferences and teacher feedback practices". *IAFOR Journal of Language Learning*, 3(2). <https://doi.org/10.22492/ijll.3.2.02>
- Kaufman, J., and Schunn, C. (2011). "Students' perceptions about peer assessment for writing: Their origin and impact on revision work". *Instructional Science*, 39(3), 387-406.
- Kerr, P. (2020). Giving feedback to language learners. In *Cambridge Papers in ELT* (pp. 1-25). Cambridge University Press.

- Khan, Z. R. (2014). Developing a factor-model to understand the impact of factors on higher education students' likelihood to e-cheat (Doctoral dissertation). University of Wollongong. Retrieved on 10 August 2023 from <https://ro.uow.edu.au/theses/4545>
- Khan, Z. R. (2024). Ethics of artificial intelligence in academia. In S. E. Eaton (Ed.), *Second Handbook of Academic Integrity* (pp. 1551-1582). Springer. https://doi.org/10.1007/978-981-287-079-7_188-1
- Kilbourne, B. (1990). Constructive feedback: Learning the art: The story of Oliver and Taylor. Ontario Institute for Studies in Education Press and Brookline Books.
- Kroll, B. (2001). Considerations for teaching an ESL/EFL writing course. In M. Celce Murcia (Ed.), *Teaching English as a second or foreign language* (3rd ed., pp. 219-232). Heinle and Heinle.
- Kumar, R., Eaton, S. E., Mindzak, M., and Morrison, R. (2024). Academic integrity and artificial intelligence: An overview. In S. E. Eaton (Ed.), *Second Handbook of Academic Integrity* (pp. 1583-1596). Springer. https://doi.org/10.1007/978-3-031-54144-5_153
- Kumaravadivelu, B. (2006). *Understanding language teaching: From method to postmethod*. Lawrence Erlbaum Associates.
- Lancaster, T. (2023). "Artificial intelligence, text generation tools and ChatGPT – Does digital watermarking offer a solution?". *International Journal of Educational Integrity*, 19, 10. <https://doi.org/10.1007/s40979-023-00131-6>
- Law, L. (2024). "Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review". *Computers and Education Open*, 6(2024). <https://doi.org/10.1016/j.caeo.2024.100174>
- Lee, I. (2008). "Student reactions to teacher feedback in two Hong Kong secondary classrooms". *Journal of Second Language Writing*, 17(3), 144-164. <https://doi.org/10.1016/j.jslw.2007.12.001>
- Lee, I. (2017). *Classroom writing assessment and feedback in L2 school contexts*. Springer. <https://doi.org/10.1007/978-981-10-3924-9>
- Lee, I. (2020). *Feedback in L2 writing classrooms*. TEFLIN Publication.
- Li, L., Liu, X., and Steckelberg, A. L. (2010). "Assessor or assessee: How student learning improves by giving and receiving peer feedback". *British Journal of Educational Technology*, 41(3), 525-536. <https://doi.org/10.1111/j.1467-8535.2009.00968.x>

- Li, L., Liu, X., and Zhou, Y. (2012). "Give and take: A re-analysis of assessor and assessee's roles in technology-facilitated peer assessment". *British Journal of Educational Technology*, 43(3), 376-384. <https://doi.org/10.1111/j.1467-8535.2011.01180.x>
- Lightbown, P., and Spada, N. (2006). *How languages are learned* (3rd ed.). Oxford University Press.
- Liu, F., Du, J., Zhou, D. Q., and Huang, B. (2020). "Exploiting the potential of peer feedback: The combined use of face-to-face feedback and e-feedback in doctoral writing groups". *Assessing Writing*, 100482. <https://doi.org/10.1016/j.asw.2020.100482>
- Liu, J., and Hansen, J. G. (2002). *Peer response in second language writing classrooms* (2nd ed.). University of Michigan Press.
- Lu, J., and Law, N. (2012). "Online peer assessment: Effects of cognitive and affective feedback". *Instructional Science*, 40(2), 257-275. <https://doi.org/10.1007/s11251-011-9177-2>
- Lu, J., and Law, N. (2012). "Online peer assessment: Effects of cognitive and affective feedback". *Instructional Science*, 40(2), 257-275. <https://doi.org/10.1007/s11251-011-9177-2>
- Lundstrom, K., and Baker, W. (2009). "To give is better than to receive: The benefits of peer review to the reviewer's own writing". *Journal of Second Language Writing*, 18(1), 30-43. <https://doi.org/10.1016/j.jslw.2008.06.002>
- Martin, J. R. (1997). Analyzing genre: Functional parameters. In F. Christie and J. R. Martin (Eds.), *Genres and institutions* (pp. 3-39). Cassell.
- Matsuda, P. K. (2003). Second language writing in the twentieth century: A situated historical perspective. In B. Kroll (Ed.), *Exploring the dynamics of second language writing* (pp. 15-34). Cambridge University Press.
- McCarthy, J. (1971). "Recursive functions of symbolic expressions and their computation by machine, part I". *Communications of the ACM*, 3(4), 184-195.
- McDonald, N., Johri, A., Ali, A., and Hingle, A. (2024). "Generative artificial intelligence in higher education: Evidence from an analysis of institutional policies and guidelines". ArXiv preprint. <https://doi.org/10.48550/arXiv.2402.01659>
- McGuire, A., Qureshi, W., and Saad, M. (2024). "A constructivist model for leveraging GenAI tools for individualized, peer-simulated feedback on student writing".

- International Journal of Technology in Education (IJTE)*, 7(2), 326-352.
<https://doi.org/10.46328/ijte.639>
- Meyer, K. A. (2006). "Student perceptions of face-to-face and online discussions: The advantage goes to...". *Journal of Asynchronous Learning Networks*, 10(4), 83-97.
<https://doi.org/10.24059/olj.v11i4.1715>
- Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., and Gerardou, F. S. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *Education Sciences*, 13(9), 856.
<https://doi.org/10.3390/educsci13090856>
- Min, H. (2005). "Training students to become successful peer reviewers". *System*, 33(2), 293-308. <https://doi.org/10.1016/j.system.2004.11.003>
- Min, H.-T. (2016). "Effect of teacher modelling and feedback on EFL students' peer review skills in peer review training". *Journal of Second Language Writing*, 31, 43-57.
<https://doi.org/10.1016/j.jslw.2016.01.004>
- Minsky, M. (1969). *Perceptrons: An introduction to computational geometry*. MIT Press.
- Montgomery, J. L., and Baker, W. (2007). Teacher-written feedback: Student perceptions, teacher self-assessment, and actual teacher performance. *Journal of Second Language Writing*, 16(2), 82-99. <https://doi.org/10.1016/j.jslw.2007.04.002>
- Morris, F. (2005). "Child-to-child interaction and corrective feedback in a computer-mediated L2 class". *Language Learning and Technology*, 9(1), 29-45.
<http://dx.doi.org/10125/44007>
- Moser, A. (2020). *Written corrective feedback: The role of learner engagement—A practical approach*. Springer Nature.
- Moya, B. A., Eaton, S. E., Pethrick, H., Hayden, K. A., Brennan, R., Wiens, J., and McDermott, B. (2023). "Academic integrity and artificial intelligence in higher education contexts: A rapid scoping review". *Canadian Perspectives on Academic Integrity*, 5(2). <https://doi.org/10.11575/cpai.v5i2.78123>
- Myles, J. (2002). "Second language writing and research: The writing process and error analysis in student texts". *Teaching English as a Second or Foreign Language*, 6(2), 1-19.
- Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. J. G. van Merriënboer, J. M. Spector, M. D. Merrill, and M. P. Driscoll (Eds.), *Handbook of*

- research on educational communications and technology* (3rd ed., pp. 125-144). Lawrence Erlbaum.
- Nelson, M. M., and Schunn, C. D. (2009). "The nature of feedback: How different types of peer feedback affect writing performance". *Instructional Science*, 37, 375-401. <https://doi.org/10.1007/s11251-008-9053-x>
- Newell, A., and Simon, H. A. (1975). "Computer science as empirical inquiry: Symbols and search". *Communications of the ACM*, 19(3), 113-126.
- Nguyen, A., Hong, Y., Dang, B., and Huang, X. (2024). "Human-AI collaboration patterns in AI-assisted academic writing". *Studies in Higher Education*, 49(5), 847-864. <https://doi.org/10.1080/03075079.2024.2323593>
- Nunan, D. (1999). *Second language teaching and learning*. Heinle and Heinle.
- Nunan, D. (2015). *Teaching English to speakers of other languages: An introduction*. Routledge.
- Nuthall, G. (2007). *The hidden lives of learners*. NZCER Press.
- Oliver, R., and Adams, R. (2021). Oral corrective feedback. In H. Nassaji and E. Kartchava (Eds.), *The Cambridge handbook of corrective feedback in second language learning and teaching* (pp. 187-206). Cambridge University Press.
- Paltridge, B., Harbon, L., Hirsh, D., Shen, H., Stevenson, M., Phakiti, A., and Woodrow, L. (2009). *Teaching academic writing: An introduction for teachers of second language writers* (Vol. 10). University of Michigan Press.
- Patchan, M. M., Charney, D., and Schunn, C. D. (2009). "A validation study of students' end comments: Comparing comments by students, a writing instructor, and a content instructor". *Journal of Writing Research*, 1(2), 124-152. <https://doi.org/10.17239/jowr-2009.01.02.2>
- Pearl, J. (2011). *Causality: Models, reasoning, and inference*. Cambridge University Press.
- Pinaya, W. H., Graham, M. S., Kerfoot, E., Tudosiu, P. D., Dafflon, J., Fernandez, V., and Cardoso, M. J. (2023). "Generative AI for medical imaging: Extending the MONAI framework". *ArXiv preprint*. <https://doi.org/10.48550/arXiv.2307.15208>
- Prensky, M. (2001). "Digital natives, digital immigrants: Part 1". *On the Horizon*, 9(5), 1-6. <https://doi.org/10.1108/10748120110424816>
- Qi, C., Tang, Y., and Lei, Y. (2024). Does feedback from ChatGPT help? Investigating the effect of feedback from both teacher and ChatGPT on students' learning outcomes.

- In *AMCIS 2024 Proceedings*. Retrieved on 10 August 2023 from https://aisel.aisnet.org/amcis2024/is_education/is_education/29
- Razi, S. (2011). *Advanced reading and writing skills in ELT: APA style handbook*. Nobel.
- Razi, S. (2015). Development of a rubric to assess academic writing incorporating plagiarism detectors. *SAGE Open*, 5(2), 1-13. <https://doi.org/10.1177/2158244015590162>
- Razi, S. (2016). "Are review skills and academic writing skills related? An exploratory analysis via multi-source feedback tools". *International Journal of Progressive Education*, 12(1), 117-127. Retrieved on 10 August 2023 from <https://dergipark.org.tr/en/pub/ijpe/issue/26318/277358>
- Razi, S. (2017). Anonymous multi-mediated writing model: Peer feedback exchange in EAP. In D. Köksal (Ed.), *Researching ELT: Classroom methodology and beyond* (pp. 29-42). Peter Lang.
- Razi, S. (2023). "Emergency remote teaching adaptation of the anonymous multi-mediated writing model". *System*, 113, 102981. <https://doi.org/10.1016/j.system.2023.102981>
- Richards, J. C., and Schmidt, R. W. (2011). *Longman dictionary of language teaching and applied linguistics* (4th ed.). Routledge.
- Richards, J., and Rodgers, T. (2001). *Approaches and methods in language teaching* (3rd ed.). Cambridge University Press.
- Risko, E. F., and Gilbert, S. J. (2016). "Cognitive offloading". *Trends in Cognitive Sciences*, 20, 676-688. <https://doi.org/10.1016/j.tics.2016.07.002>
- Rodrigues, M., Silva, R., Borges, A. P., Franco, M., and Oliveira, C. (2024). "Artificial intelligence: Threat or asset to academic integrity? A bibliometric analysis". *Kybernetes*. <https://doi.org/10.1108/K-09-2023-1666>
- Rollinson, P. (2005). "Using peer feedback in the ESL writing class". *ELT Journal*, 59(1), 23-30. <https://doi.org/10.1093/elt/cci003>
- Ruegg, R. (2015). "The relative effects of peer and teacher feedback on improvement in EFL students' writing ability". *Linguistics and Education*, 29, 73-82. <https://doi.org/10.1016/j.linged.2014.12.001>
- Russell, S. J., and Norvig, P. (2021). *Artificial Intelligence: A modern approach* (4th ed.). Pearson.
- Sauro, S. (2009). "Computer-mediated corrective feedback and the development of L2 grammar". *Language Learning and Technology*, 13(1), 96-120. <http://dx.doi.org/10.125/44170>

- Schultz, J. M. (2000). Computers and collaborative writing in the foreign language curriculum. In M. Warschauer and R. Kern (Eds.), *Network-based language teaching: Concepts and practice* (pp. 121-150). Cambridge University Press.
- Shaw, C., Yuan, L., Brennan, D., Martin, S., Janson, N., and Fox, K. (2023, October 23). Time for class - 2023. Tyton Partners. Retrieved on 10 August 2023 from <https://tytonpartners.com/time-for-class-2023/GenAI-Update>
- Sheen, Y. (2007). "The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles". *TESOL Quarterly*, 41, 255-283. <https://doi.org/10.1002/j.1545-7249.2007.tb00059.x>
- Stake, R. E. (1995). *The art of case study research*. Sage Publications.
- Stake, R. E. (2005). Qualitative case studies. In N. K. Denzin and Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd ed., pp. 443-466). Sage Publications.
- Stake, R. E. (2006). *Multiple case study analysis*. Guilford Press.
- Storch, N. (2011). "Collaborative writing in L2 contexts: Processes, outcomes, and future directions". *Annual Review of Applied Linguistics*, 31, 275-288. <https://doi.org/10.1017/S0267190511000079>
- Sugita, Y. (2006). "The impact of teachers' comment types on students' revision". *ELT Journal*, 60(1), 34-41. <https://doi.org/10.1093/elt/cci079>
- Sullivan, M., Kelly, A., and McLaughlan, P. (2023). "ChatGPT in higher education: Considerations for academic integrity and student learning". *Journal of Applied Learning and Teaching*, 6(1), 1-10. <https://doi.org/10.37074/jalt.2023.6.1.17>
- Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.
- Taiye, M., High, C., Velander, J., Matar, K., Okmanis, R., and Milrad, M. (2024). Generative AI-enhanced academic writing: A stakeholder-centric approach for the design and development of CHAT4ISP-AI. In *Proceedings of the ACM* (pp. 74-80). <https://doi.org/10.1145/3605098.3636055>
- Tauginienė, L., Gaižauskaitė, I., Glendinning, I., Kravjar, J., Ojsteršek, M., Ribeiro, L., and Sivasubramaniam, S. (2018). Glossary for academic integrity. *ENAI Report 3G*. Retrieved on 10 August 2023 from http://www.academicintegrity.eu/wp/wp-content/uploads/2018/02/GLOSSARY_final.pdf.
- The Economist. (2024, July 16). A short history of AI. Retrieved on 10 August 2023 from <https://www.economist.com/schools-brief/2024/07/16/a-short-history-of-ai>

- Topping, K. (1998). "Peer assessment between students in colleges and universities". *Review of Educational Research*, 68(3), 249-276. <https://doi.org/10.3102/00346543068003249>
- Uzun, K. (2024). Enhancing written communication skills for academic purposes. In E. Zehir Topkaya and H. Çelik (Eds.), *English for academic purposes: Theory into practice*. Springer Nature.
- Van Wyk, M. M. (2024). "Is ChatGPT an opportunity or a threat? Preventive strategies employed by academics related to a GenAI-based LLM at a faculty of education". *Journal of Applied Learning and Teaching*, 7(1), 35-45. <https://doi.org/10.37074/jalt.2024.7.1.15>
- Vinagre, M., and Muñoz, B. (2011). "Computer-mediated corrective feedback and language accuracy in telecollaborative exchanges". *Language Learning and Technology*, 15(1), 72-103. <https://doi.org/10.10125/44238>
- Vygotsky, L. (1978). *Mind in society*. Harvard University Press.
- Weaver, M. R. (2006). "Do students value feedback? Student perceptions of tutors' written responses". *Assessment and Evaluation in Higher Education*, 31(3), 379-394. <https://doi.org/10.1080/02602930500353061>
- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., and Schmidt, D. C. (2023). A prompt pattern catalog to enhance prompt engineering with ChatGPT. *ArXiv preprint*. <https://doi.org/10.48550/arXiv.2302.11382>
- Williams, J. (2012). "The potential role(s) of writing in second language development". *Journal of Second Language Writing*, 21(4), 321-331. <https://doi.org/10.1016/j.jslw.2012.09.007>
- Wood, D., Bruner, J. S., and Ross, G. (1976). "The role of tutoring in problem solving". *Journal of Child Psychology and Psychiatry*, 17(2), 89-100. <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>
- Ye, Q., Axmed, M., Pryzant, R., and Khani, F. (2023). "Prompt engineering a prompt engineer". *ArXiv preprint*. <https://doi.org/10.48550/arXiv.2311.05661>
- Yoshida, R. (2008). "Teachers' choice and learners' preference of corrective feedback types". *Language Awareness*, 17(1), 78-93. <https://doi.org/10.2167/la429.0>
- Yusuf, A., Pervin, N., and Román-González, M. (2024). "Generative AI and the future of higher education: A threat to academic integrity or reformation? Evidence from

- multicultural perspectives”. *International Journal of Educational Technology in Higher Education*, 21, 21. <https://doi.org/10.1186/s41239-024-00453-6>
- Zhan, L. (2016). “Written teacher feedback: Student perceptions, teacher perceptions, and actual teacher performance”. *English Language Teaching*, 9(8), 73-84. <https://doi.org/10.5539/elt.v9n8p73>
- Zhang, L. J. (2013). “Second language writing as and for second language learning”. *Journal of Second Language Writing*, 22(4), 446-447. <https://doi.org/10.1016/j.jslw.2013.08.010>



APPENDICES

APPENDIX A

WRITING SKILLS I COURSE CONTENT

1

Title of the Course: <i>Writing Skills 1</i>	Course-No: <i>İNÖ-1003</i>	Semester: <i>1st Year</i> <i>1st Semester (Fall)</i>	
Course type: <i>Lecture</i>	Hours-Practice/Week: <i>2-0</i>	Credits	
		Local	ECTS
Status: <i>Compulsory</i>	Instruction: <i>English</i>	<i>2</i>	<i>2</i>
Lecturer:	<i>Prof. Dr. S</i>		

Course description:

This course aims to develop academic written skills by means of critical analysis of academic texts. Related readings will come from the relevant literature in various ways such as academic books and journals which enable learners to develop and increase vocabulary in an academic context with the hope of contributing to their writing skills. Within this perspective, this course mainly encourages learners to write academic texts. Learners will have the opportunity of benefiting from scaffolding through digital peer feedback exchange on their written assignments. Learners are expected to analyse and produce essays considering the organization, coherence, cohesion, unity, vocabulary selection, in-text citations rules including quoting, paraphrasing, and summarizing as this course aims to teach academic writing to students who need to present and defend their own points of view clearly and persuasively. Students also work to master the norms of in-text citations and attributing data from print and other sources, and to observe the methods of acknowledging the research findings of others, specifically in the APA style.

Teaching method: Lectures, extensive reading, peer-review.

Prerequisites: Advanced reading comprehension skills and proficiency in informal writing.

Teaching aids: PowerPoint presentation, Turnitin (text-matching software)

Assessment method:

There are no pen and paper tests for this course; however, it does not necessarily mean that you will pass without any efforts. During the term, you need to submit 3 written assignments at an average length of 700 words for each. For each assignment, you will receive feedback on earlier drafts of your paper. You will then and revise and resubmit your paper by taking the feedback into consideration. For your first assignment, you will receive teacher feedback whereas you will exchange peer feedback for your second assignment. For this purpose, you will receive feedback from three peers and provide feedback to three of your peers that I will assign. For your third assignment, you will benefit from artificial intelligence to receive feedback by using the prompts that I instruct. All revision assignments need to be submitted digitally both with and without tracks of changes. The process of submitting assignments and exchanging feedback will be managed through Turnitin. Hard-copy or email submissions will not be taken into consideration. I will evaluate the accuracy and the quality of the peer feedback that you provide and scores retrieved from peer review tasks will function as mid-term examination with the weight of 40% whereas scores retrieved from writing tasks will function as final examination and have the weight of 60%.

Attendance policy:

It is important that you attend and participate in class; our class meets only once a week, so missing one class represents a substantial portion of the semester. If there are special circumstances requiring you to be out of class, please email me. Attendance will be taken and recorded in all class sessions as instructed by institutional policies. Please follow institutional announcements regarding the changes in the attendance policy during emergency situations. Please do remember to come to class prepared and on time.

Academic integrity policy:

This course aims to cultivate integrity among undergraduates as any kind of educational operation should be affiliated with the principles of academic integrity. Plagiarism is a serious breach of academic trust. In academic work, our words and ideas are the value of our work, so turning in someone else's work as if it were your own is a form of theft. When you use someone else's words and ideas, whether it is the work of a famous author or a fellow student, without crediting the source or authorship of those words and ideas, you are plagiarizing. So, here is the bottom line: original work only, credit to ideas, writing, or words from someone other than you. Either accidental or intentional, there is no tolerance for any type of academic misconduct. You need to submit all your assignments through Turnitin and as the lecturer I will determine the originality of your assignments by checking them against plagiarism with regards to the similarity reports. Using AI to generate a written assignment is an academic misconduct. Please make sure that you follow the instructions very carefully relating to receiving feedback from AI tools; otherwise, you may be accused of academic misconduct. In the case of plagiarism, collusion, contract cheating, unethical use of AI, you will receive a "0" for the assignment. Please check our academic integrity policy for further details.

Registration for the examination:

To submit your assignments and to review your peer's assignments through Turnitin, you need to enrol in Turnitin virtual class.

Assignment requirements:

All assignment must be on A-4 paper with 2.5cm margins on all sides, and be written with double spacing using Times New Roman 12 font. All must be turned in on time only through Turnitin. Late assignments will not be taken into consideration. Do NOT write your name or faculty number on your assignments to enable anonymous peer-review.

Main reading:

- Razi, S. (2011). *Advanced reading and writing skills in ELT: APA style handbook*. Ankara: Nobel.
- Razi, S. (2015a). Cross-checked problems in undergraduate academic writing. K. Dikilitas, R. Smith, & W. Trotman (Eds.), *Teacher-researchers in action* (pp. 147-161). Kent, UK: IATEFL.
- Razi, S. (2015b). Development of a rubric to assess academic writing incorporating plagiarism detectors. *SAGE Open*, 5(2), 1-13. DOI: 10.1177/2158244015590162
- Razi, S. (2016a). Are review skills and academic writing skills related? An exploratory analysis via multi-source feedback tools. *International Journal of Progressive Education*, 12(1), 117-127.
- Razi, S. (2016b). Open and anonymous peer review in a digital online environment compared in academic writing context. In C. Goria, O. Speicher & S. Stollhans (Eds.), *Innovative language teaching and learning at university: Enhancing participation and collaboration* (pp. 49-56). Dublin, UK: Research-publishing.net.
- Razi, S. (2016c). Undergraduates' perceptions of anonymous and open digital peer feedback in academic writing. In P. Sappapan (Ed.), *Foreign language learning and teaching 4* (pp. 78-89). Bangkok: Thammasat University.
- Razi, S. (2017a). Anonymous multi-mediated writing model: Peer feedback exchange in EAP. In D. Köksal (Ed.), *Researching ELT: Classroom methodology and beyond* (pp. 29-42). Frankfurt: Peter Lang.
- Razi, S. (2017b). *Benefiting from anonymous multi-mediated writing model to avoid plagiarism*. Paper presented at the European Network for Academic Integrity 1st Joint Staff Training Event, 13 September, University of Maribor, Slovenia.
- Razi, S. (2017c). *How to benefit from similarity reports?* Paper presented at the European Network for Academic Integrity 1st Joint Staff Training Event, 14 September, Ministry of Education, Science and Sports, Ljubljana, Slovenia.
- Razi, S. (2018). *Fallacies in the interpretation of text-matching software similarity reports*. Paper presented to be at the 4th Plagiarism across Europe and Beyond Conference, Fostering academic integrity through consistent and transparent approaches, 9-11 May 2018, Ephesus, Turkey.
- Razi, S. (2022). Academic integrity and referencing: Referencing. In A. Boney (Ed.), *SAGE skills: Student success*. SAGE. <https://sk.sagepub.com/skills/student-success/air/referencing> / <https://dx.doi.org/10.4135/9781071893630>

Registration for the examination:

To submit your assignments and to review your peer's assignments through Turnitin, you need to enrol in Turnitin virtual class.

Assignment requirements:

All assignment must be on A-4 paper with 2.5cm margins on all sides, and be written with double spacing using Times New Roman 12 font. All must be turned in on time only through Turnitin. Late assignments will not be taken into consideration. Do NOT write your name or faculty number on your assignments to enable anonymous peer-review.

Main reading:

- Razi, S. (2011). *Advanced reading and writing skills in ELT: APA style handbook*. Ankara: Nobel.
- Razi, S. (2015a). Cross-checked problems in undergraduate academic writing. K. Dikilitaş, R. Smith, & W. Trotman (Eds.), *Teacher-researchers in action* (pp. 147-161). Kent, UK: IATEFL.
- Razi, S. (2015b). Development of a rubric to assess academic writing incorporating plagiarism detectors. *SAGE Open*, 5(2), 1-13. DOI: 10.1177/2158244015590162
- Razi, S. (2016a). Are review skills and academic writing skills related? An exploratory analysis via multi-source feedback tools. *International Journal of Progressive Education*, 12(1), 117-127.
- Razi, S. (2016b). Open and anonymous peer review in a digital online environment compared in academic writing context. In C. Gorla, O. Speicher & S. Stollhans (Eds.), *Innovative language teaching and learning at university: Enhancing participation and collaboration* (pp. 49-56). Dublin, UK: Research-publishing.net.
- Razi, S. (2016c). Undergraduates' perceptions of anonymous and open digital peer feedback in academic writing. In P. Sappapan (Ed.), *Foreign language learning and teaching 4* (pp. 78-89). Bangkok: Thammasat University.
- Razi, S. (2017a). Anonymous multi-mediated writing model: Peer feedback exchange in EAP. In D. Köksal (Ed.), *Researching ELT: Classroom methodology and beyond* (pp. 29-42). Frankfurt: Peter Lang.
- Razi, S. (2017b). *Benefiting from anonymous multi-mediated writing model to avoid plagiarism*. Paper presented at the European Network for Academic Integrity 1st Joint Staff Training Event, 13 September, University of Maribor, Slovenia.
- Razi, S. (2017c). *How to benefit from similarity reports?* Paper presented at the European Network for Academic Integrity 1st Joint Staff Training Event, 14 September, Ministry of Education, Science and Sports, Ljubljana, Slovenia.
- Razi, S. (2018). *Fallacies in the interpretation of text-matching software similarity reports*. Paper presented to be at the 4th Plagiarism across Europe and Beyond Conference, Fostering academic integrity through consistent and transparent approaches, 9-11 May 2018, Ephesus, Turkey.
- Razi, S. (2022). Academic integrity and referencing: Referencing. In A. Boney (Ed.), *SAGE skills: Student success*. SAGE. <https://sk.sagepub.com/skills/student-success/air/referencing> / <https://dx.doi.org/10.4135/9781071893630>

Razi, S. (2023). Emergency remote teaching adaptation of the anonymous multi-mediated writing model. *System*, 113, 102981.

<https://doi.org/10.1016/j.system.2023.102981>

Supplementary reading:

American Psychological Association (APA). (2020). *Publication manual of the American Psychological Association* (7th ed.). APA.

Bailey, S. (2011). *Academic writing: A handbook for international students*. New York, NY: Routledge.

Cumming, A. (2006). *Goals for academic writing: ESL students and their instructors*. Amsterdam: John Benjamins.

Hashimoto, I. Y., Kroll, B. M., & Schafer, J. C. (1982). *Strategies for academic writing: A guide for college students*. Ann Arbor, MI: The University of Michigan Press.

Oshima, A. (1997). *Introduction to academic writing*. White Plains: Longman.

Peachy, N. (2017). *Digital tools for teachers*. Peachey Publications.

Reid, J. M. (2000). *The process of composition* (3rd ed.). Englewood Cliffs: Prentice Hall Regents.

Zemach, D. E., & Rumisek, L. A. (2003). *Academic writing: From paragraph to essay*. Oxford: Macmillan.

Williams, K., & Davis, M. (2017). *Referencing & understanding plagiarism* (2nd ed.). London: Palgrave.

Week / Date	Discussion Topics	Main Reading	Supplementary Reading	Assignments
Week 1 2-6.10.2023	Introducing course content.	-----	-----	-----
Week 2 9-13.10.2023	Introduction to academic writing.	Razi (2011) [Chapter 1]	APA (2020) Cumming (2006) Oshima (1997)	-----
Week 3 16-20.10.2023	Basic steps in process writing. International Day of Action for Academic Integrity.	Razi (2015a, 2015b) Razi (2023)	Bailey (2011) Reid (2000)	-----
Week 4 23-27.10.2023	Introduction to in-text citations.	Razi (2011) [Chapter 8]	APA (2020)	Assignment 1 (First draft)
Week 5 30.10-3.11.2023	Presenting in-text citations.	Razi (2011) [Chapter 9]	APA (2020) Hashimoto et al. (1982) Zemach & Rumisek (2003)	Assignment 1 (Teacher feedback)
Week 6 6-10.11.2023	Familiarization of academic integrity: Avoiding plagiarism and contract cheating.	Razi (2017b)	Bailey (2011) Williams & Davis (2017, pp. 67-94)	Assignment 1 (Revision)
Week 7 13-17.11.2023	Interpreting text-matching software similarity reports.	Razi (2017c, 2018)	-----	-----
Week 8 20-24.11.2023	Midterm exam week.	-----	-----	Assignment 2 (First draft)
Week 9 27.11-1.12.2023	Details with in-text citations.	Razi (2011) [Chapter 11]	APA (2020)	Assignment 2 (Peer feedback)
Week 10 4-8.12.2023	Exchanging effective peer feedback.	Razi (2016a, 2016b, 2016c, 2017a)	-----	Assignment 2 (Revision)
Week 11 11-15.12.2023	Practising in-text citations.	Razi (2011) [Chapter 10]	APA (2020)	-----
Week 12 18-22.12.2023	Prompt writing in AI tools	-----	-----	Assignment 3 (First draft)
Week 13 25-29.12.2023	Practising prompt writing in AI tools	-----	-----	Assignment 3 (AI feedback)
Week 14 1-5.01.2024	Connecting ideas: Unity and coherence.	Razi (2011) [Chapter 6]	Bailey (2011)	Assignment 3 (Revision)
8-19.01.2024	Final Exam Week	-----	-----	-----

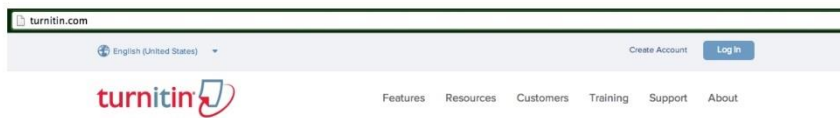
Turnitin Assignment Submission Guidelines

Enrolling in a class

You must enroll in your class before you can begin using Turnitin. To do this, use the class ID and enrollment password above. To create your account, please visit <http://turnitin.com/> and follow the instructions below.

How to create a new user profile and enrol in your class

1. Click the “**Create Account**” link to the left of the Sign in button on the Turnitin homepage.



2. Click on the **student link** from the Create a New Turnitin Account screen.

 A screenshot of the 'Create a User Profile' screen. The title 'Create a User Profile' is at the top. Below it, the section 'Have You Ever Used Turnitin?' is followed by a paragraph: 'If you've used Turnitin before, you can use the same email and password to log in. You can keep all your papers and grades together, even if you're now in a different class or a different school!'. There are two input fields: 'Email address' and 'Password (Login to Turnitin)'. Below the password field is a link: 'Forgot your password? Click here.' The next section is 'Create a New Account', followed by the instruction: 'Please select whether you will be using the service as an instructor or a student.' There are three radio button options: 'Student', 'Instructor', and 'Teaching assistant'. At the bottom, there is a 'Login' button.

3. Enter your **class ID** and **enrollment password** in the appropriate fields.

Create a New Student Account

Class ID Information

All students must be enrolled in an active class. To enroll in a class, please enter the class ID number and class enrollment password that you were given by your instructor.

Please note that the password and pincode are case-sensitive. If you do not have this information, or the information you are entering appears to be incorrect, please contact your instructor.

Class ID

Class enrollment password

4. Enter your **first and last names** and the **e-mail address** you would like associated with your Turnitin account. This will be your user name.

User Information

Your first name

Your last name

Display names as

☒ First name (Space) Last name (example: John Smith)
☐ Last name (Space) First name (example: Smith John)
☐ Last name(No space)First name (example: SmithJohn)

Email address

5. Enter and re-enter the **password** you would like to use as your Turnitin login password.

Password and Security

Please keep in mind that your password is case sensitive (for example, paSS123 would be different than pass123), must be 6-12 characters long, and contain at least one letter and one number (for added security).

Enter an email address and password you can easily remember. Please write down your email address and password for future reference.

Enter your password

Confirm your password

APPENDIX B

FUNDAMENTALS OF ACADEMIC WRITING RUBRIC

Questions	Rating (Out of 40 points)		
Introduction	0 point	1 point	2 points
1. <i>Topic relevance</i>	Irrelevant topic.	The topic is somewhat relevant.	The topic is extremely relevant.
2. <i>Topic introduction</i>	No introduction of the topic or the problem situation.	Improper introduction of the topic and problem situation.	Proper introduction of the topic and problem situation.
3. <i>Thesis statement</i>	No introduction of the thesis statement.	Improper introduction of the thesis statement.	Proper introduction of the thesis statement.
Citation	0 point	1 point	2 points
4. <i>Citations (Paraphrases & Summaries)</i>	Citations are largely inconsistent and incorrect.	Citations are mostly documented properly.	Citations are properly documented.
5. <i>Direct quotations</i>	Quotations are largely inconsistent and incorrect.	Quotations are mostly documented properly.	Quotations are properly documented.
6. <i>Citation needed</i>	Most expressions lack citations.	Few expressions require citations.	No expression lacks citation.
Academic Writing	0 point	1 point	2 points
7. <i>Formality level</i>	Expressions are largely colloquial. Extensive use of informal words.	Some expressions are colloquial. Informal vocabulary choice is visible.	Expressions are non-colloquial. Academic vocabulary choice is dominant.
8. <i>Passive voice</i>	No use of passive expressions.	Redundant use of passive expressions decreases readability.	Passive expressions are used only when necessary to contribute to academic tone.
9. <i>Complexity</i>	Ideas are presented in short and choppy sentences.	Attempts to use conjunctions for sophisticated connections.	Proper use of conjunctions for sophisticated connections.
10. <i>Abbreviations</i>	Incorrect use of abbreviations.	Inconsistent use of abbreviations.	Proper use of abbreviations.

<i>Idea Presentation</i>	<i>0 point</i>	<i>1 point</i>	<i>2 points</i>
<i>11. Structure of the essay</i>	Essay lacks focus as the overall structure is disorganized.	Focus of the essay is improper as the overall structure has some flaws.	Introduction is followed by well-developed discussion paragraphs and a relevant conclusion section.
<i>12. Main idea and thesis statement relevance</i>	Main ideas largely do not directly contribute to the thesis statement.	Some main ideas lack direct contribution to the thesis statement.	All main ideas in directly contribute to the thesis statement.
<i>13. Conclusions and main idea relevance</i>	Concluding remarks largely are not relevant to the thesis statement.	Some concluding statements lack relevance to the thesis statement.	All concluding remarks are relevant to the thesis statement.
<i>14. Paragraph development</i>	Ill-developed paragraphs with ideas introduced unclearly without exploration or relevant conclusion.	Attempts to develop paragraphs with some flaws relating to introduction of main idea, discussion by blending other sources, and relevant concluding remark.	Well-developed paragraphs considering introduction of main idea, discussion by blending other sources, relevant concluding remark.
<i>15. Flow of ideas including unity and coherence</i>	Ideas are not linked properly by transitions resulting in disunified and incoherent paragraphs.	Ideas are mostly linked properly by transitions in almost unified and coherent paragraphs.	Ideas are linked properly by transitions in unified and coherent paragraphs.
<i>16. Concluding remarks</i>	Concluding remarks are largely ineffective and irrelevant to the ideas discussed.	Concluding remarks are largely effective and relevant to the ideas discussed.	All concluding remarks are effective and relevant to the ideas discussed.
<i>Mechanics</i>	<i>0 point</i>	<i>1 point</i>	<i>2 points</i>
<i>17. Grammar</i>	There are several grammar errors preventing	There are a few grammar errors causing	There are no grammar errors.

	readability of the essay.	comprehension problems.	
<i>18. Spelling</i>	There are several spelling errors preventing readability of the essay.	There are a few spelling errors causing comprehension problems.	There are no spelling errors.
<i>19. Punctuation</i>	There are several punctuation errors preventing readability of the essay.	There are a few punctuation errors decreasing reader's enjoyment of the essay.	There are no punctuation errors.
<i>20. Short forms</i>	There are several contractions preventing readability of the essay.	There are a few contractions causing comprehension problems.	There are no contractions full words are used instead.

APPENDIX C

DEVELOPED PROMPTS

<i>Introduction</i>	
<i>1. Topic relevance</i>	Please evaluate the alignment of this essay with the course instructor's topic: 'The use of L1 in teaching English to young learners as a foreign language.' Analyze the level of relevance or irrelevance of the article to this subject and substantiate your assessment with numbered feedback, supported by evidence and examples from the text. Additionally, assign a relevance score (integer) to the essay according to the following criteria: 0 for irrelevant, 1 for somewhat relevant, and 2 for highly relevant.
<i>2. Topic introduction</i>	Evaluate the introductory section of this essay concerning its ability to engage the reader, provide a clear overview of the topic, and establish the context for the discussion. Provide specific feedback with numbered points, substantiated by evidence and examples from the text. Additionally, rate the topic introduction of the whole essay using a single score (integer) on a scale of 0 to 2 based on the following criteria: 0 for no introduction of the topic or problem, 1 for an inadequate introduction, and 2 for an effective introduction.
<i>3. Thesis statement</i>	Assess the essay's thesis statement, considering its clarity (how easily it can be understood), specificity (focus on a particular aspect), argumentative nature (clear stance presented), and scope (appropriate breadth). Provide specific feedback with numbered points, supported by evidence and examples from the text. Additionally, assign a single score (whole number) to the thesis statement of the essay on a scale of 0 to 2 according to the following criteria: 0 for no thesis statement introduction, 1 for an inadequately introduced thesis statement, and 2 for a well-introduced thesis statement.
<i>Citation</i>	

4. <i>Citations (Paraphrases & Summaries)</i>	Evaluate the essay's citations in accordance with APA 7th edition, including an assessment of paraphrases and summaries. Offer specific feedback using numbered points and provide evidence and examples from the text to support your assessments. Furthermore, allocate a single score (whole number) on a scale of 0 to 2 based on the following criteria: 0 for "Citations are largely inconsistent and incorrect," 1 for "Citations are mostly documented properly," and 2 for "Citations are properly documented."
5. <i>Direct quotations</i>	Please evaluate the essay's use of direct quotations following the guidelines of APA 7th edition. Offer specific feedback using numbered points and provide evidence and examples from the text to

	support your assessments. Additionally, assign a single score (whole number) on a scale of 0 to 2 based on the following criteria: 0 for "Quotations are largely inconsistent and incorrect." 1 for "Quotations are mostly documented properly." 2 for "Quotations are properly documented."
6. <i>Citation needed</i>	Evaluate the uncited sentences within the essay to determine if they necessitate citations, taking into account the distinction between common and uncommon knowledge. Offer specific feedback using numbered points and support your assessments with evidence and examples derived from the text. Additionally, assign a single score (whole number) to assess the essay on a scale of 0 to 2 according to the following criteria: 0 for "Most expressions lack citations." 1 for "A few expressions require citations." 2 for "No expressions lack citations."
Academic Writing	
7. <i>Formality level</i>	Evaluate the formality level of this academic essay. Provide specific feedback using numbered points, and substantiate your assessments with evidence and examples from the text. Additionally, assign a single score (whole number) to assess the overall formality level of the entire essay on a scale of 0 to 2, based on the following criteria: 0 for "Expressions are largely colloquial with extensive use of informal words," 1 for "Some expressions are colloquial, and informal vocabulary choice is evident," and 2 for "Expressions are non-colloquial, with a dominant use of academic vocabulary."
8. <i>Passive voice</i>	In the context of an academic essay evaluation, assess the essay's utilization of passive voice. Provide specific feedback with numbered points and substantiate your assessments with textual evidence and examples. Furthermore, assign a single numerical score, ranging from 0 to 2, to gauge the essay's overall use of passive voice. Employ the following criteria for scoring: 0 for "No utilization of passive expressions," 1 for "Redundant use of passive expressions diminishing readability," and 2 for "Passive expressions employed judiciously, contributing to an academic tone."
9. <i>Complexity</i>	Please evaluate the complexity of this academic essay using the following criteria. Your feedback should be specific, presented in numbered points, and supported by evidence and examples from the text. Additionally, assign an overall complexity score on a scale of 0 to 2 based on the criteria: 0 for "Ideas are presented in short and choppy sentences," 1 for "Attempts to use conjunctions for

	sophisticated connections," and 2 for "Proper use of conjunctions for sophisticated connections."
<i>10. Abbreviations</i>	Evaluate the usage of abbreviations in the given academic essay with careful consideration. Abbreviations can serve as valuable tools for simplifying complex or lengthy terms, particularly in scientific, technical, or specialized contexts. However, their deployment should enhance the clarity and readability of the text, not hinder it. To begin the evaluation, please confirm whether this essay contains any abbreviations. If no abbreviations are present, kindly indicate so. If the essay does include abbreviations, kindly assess their usage based on the following criteria. Additionally, provide a single score on a scale from 0 to 2 to gauge the overall application of abbreviations throughout the essay. Your evaluation should incorporate specific feedback, structured with numbered points, and supported by evidence and examples from the text. The scoring scale is as follows: 0 for "Incorrect use of abbreviations," 1 for "Inconsistent use of abbreviations," and 2 for "Proper use of abbreviations."
Idea Presentation	
<i>11. Structure of the essay</i>	Evaluate the structure of the essay in the given academic essay with careful consideration. In your assessment of the structure of this academic essay, please consider the following criteria: the essay's focus, organization, the quality of its introduction, development of discussion paragraphs, and relevance of the conclusion section. Additionally, provide a single score on a scale from 0 to 2 to gauge the overall structure of the essay. Your evaluation should incorporate specific feedback, structured with numbered points, and supported by evidence and examples from the text. The scoring scale is as follows: 0 for " Essay lacks focus as the overall structure is disorganized.", 1 for " Focus of the essay is improper as the overall structure has some flaws.", and 2 for " Introduction is followed by well-developed discussion paragraphs and a relevant conclusion section.".
<i>12. Main idea and thesis statement relevance</i>	Assess the main idea and relevance of the thesis statement in the provided academic essay. Furthermore, assign a single score to the entire essay on a scale from 0 to 2 to measure the alignment of the main idea with the thesis statement. Your evaluation should incorporate specific feedback, organized in numbered points, and substantiated with evidence and examples from the text. The scoring criteria are as follows: 0 for

	"The main ideas largely do not directly support the thesis statement," 1 for "Some main ideas lack direct support for the thesis statement," and 2 for "All main ideas directly contribute to the thesis statement."
<i>13. Conclusions and main idea relevance</i>	Assess the relevance of conclusions and their alignment with the main idea in the provided academic essay. Additionally, assign a single score to the entire essay on a scale from 0 to 2 to measure the conclusions and main idea relevance. Your evaluation should incorporate specific feedback, organized in numbered points, and substantiated with evidence and examples from the text. The scoring criteria are as follows: 0 for "Concluding remarks largely are not relevant to the thesis statement," 1 for "Some concluding statements lack relevance to the thesis statement," and 2 for "All concluding remarks are relevant to the thesis statement."
<i>14. Paragraph development</i>	Could you please provide feedback on the essay's paragraph development, specifically assessing the clarity of introduced ideas, the depth of exploration, and the relevance of conclusions, considering the spectrum from ill-developed paragraphs to attempts with flaws and ultimately to well-developed paragraphs with effective integration of main ideas, external sources, and concluding remarks? Additionally, assign a single score to the entire essay on a scale from 0 to 2 to measure the essay's paragraph development. Your evaluation should incorporate specific feedback, organized in numbered points, and substantiated with evidence and examples from the text. The scoring criteria are as follows: 0 for "Ill-developed paragraphs with ideas introduced unclearly without exploration or relevant conclusion.", 1 for "Attempts to develop paragraphs with some flaws relating to introduction of main idea, discussion by blending other sources, and relevant concluding remark.", and 2 for "Well-developed paragraphs considering introduction of main idea, discussion by blending other sources, relevant concluding remark."
<i>15. Flow of ideas including unity and coherence</i>	Could you please evaluate the essay's flow of ideas, giving specific feedback on the use of transitions and the overall unity and coherence of the paragraphs? Your assessment should focus on whether the ideas are inadequately linked, adequately linked but somewhat unified, or effectively linked in unified and coherent paragraphs. Additionally, assign a single score to the entire essay on a scale from 0 to 2, gauging the flow

	of ideas. Your evaluation should include numbered points with evidence and examples from the text to support your feedback. The scoring criteria are as follows: 0 for "Ideas are not linked properly by transitions resulting in disunified and incoherent paragraphs," 1 for "Ideas are mostly linked properly by transitions in almost unified and coherent paragraphs," and 2 for "Ideas are linked properly by transitions in unified and coherent paragraphs."
16. <i>Concluding remarks</i>	Could you please provide feedback on the essay's concluding remarks? Additionally, assign a single score to the entire essay on a scale from 0 to 2 to measure the essay's concluding remarks. Your evaluation should incorporate specific feedback, organized in numbered points, and substantiated with evidence and examples from the text. The scoring criteria are as follows: 0 for "Concluding remarks are largely ineffective and irrelevant to the ideas discussed.", 1 for "Concluding remarks are largely effective and relevant to the ideas discussed.", and 2 for "All concluding remarks are effective and relevant to the ideas discussed."
Mechanics	
17. <i>Grammar</i>	Can you kindly offer feedback on the essay's grammar, focusing on Grammar Rules and Conventions and Sentence Fragments? Additionally, please assign a single score (integer) to the overall essay on a scale from 0 to 2, assessing the essay's grammar. Your evaluation should encompass specific feedback, organized in numbered points, substantiated with evidence and examples from the text. The scoring criteria are as follows: 0 for "There are several grammar errors preventing readability of the essay," 1 for "There are a few grammar errors causing comprehension problems," and 2 for "There are no grammar errors."
18. <i>Spelling</i>	Can you kindly offer feedback on the essay's spelling? Furthermore, please assign a single score (integer) to the overall essay on a scale from 0 to 2, evaluating the essay's spelling. Your assessment should include specific feedback, organized in numbered points, and supported with evidence and examples from the text. The scoring criteria are as follows: 0 for "There are several spelling errors preventing readability of the essay," 1 for "There are a few spelling errors causing comprehension problems," and 2 for "There are no spelling errors."
19. <i>Punctuation</i>	Can you kindly offer feedback on the essay's punctuation? Furthermore, please assign a single score (integer) to the overall essay on a scale from 0 to 2, evaluating the essay's punctuation. Your

	assessment should include specific feedback, organized in numbered points, and supported with evidence and examples from the text. The scoring criteria are as follows: 0 for "There are several punctuation errors preventing readability of the essay." 1 for "There are a few punctuation errors decreasing reader's enjoyment of the essay." and 2 for "There are no punctuation errors."
20. <i>Short forms</i>	Contractions refer to the shortening of words by combining two words into one, typically using an apostrophe. For example, "can't" is a contraction of "cannot," and "it's" is a contraction of "it is" or "it has." In the context of academic writing, the use of contractions is generally discouraged. Academic essays often follow a formal writing style, and contractions can make the writing seem less formal. Please evaluate the essay's use of contractions and assign a single score (integer) on a scale from 0 to 2 to measure the essay's contractions. Your feedback should cover all relevant information provided, and the scoring criteria are as follows: 0 for "There are several contractions preventing readability of the essay," 1 for "There are a few contractions causing comprehension problems," and 2 for "There are no contractions; full words are used instead." Additionally, please ensure that your evaluation includes specific feedback, organized in numbered points, and supported with evidence and examples from the text.

APPENDIX D

THE FACULTY OF EDUCATION COMMITTEE APPROVAL

Sayı : E-68203582-605.01-2300261931
Konu : Veri Toplama İsteminiz

26.10.2023

Sayın Burcu ÖZGE RAZI

Prof. Dr. Aysun YAVUZ'un danışmanlığında yürüttüğünüz "Ethical Implementation of AI in the Process of Academic Writing (Yapay Zekanın Akademik Yazım Sürecinde Etik Kullanımı)" başlıklı yüksek lisans teziniz kapsamında Fakültemiz İngilizce Öğretmenliği programı 1. sınıf öğrencilerinden, 2023-2024 Akademik Yılı Güz Yarıyılında aldıkları "Yazma Becerileri I" dersinde veri toplama isteminiz Fakültemiz Bilimsel Araştırmaları Değerlendirme Kurulu tarafından incelenmiş ve uygun görülmüştür. Bilgilerinizi ve gereğini rica ederim.



APPENDIX E

THE ETHICS COMMITTEE APPROVAL



T.C.
ÇANAKKALE ONSEKİZ MART ÜNİVERSİTESİ REKTÖRLÜĞÜ
Lisansüstü Eğitim Enstitüsü Etik Kurulu
Bilimsel Araştırma ve Yayın Etiği



Sayı : E-84026528-050.01.04-2300289440
Konu : Başvuru İncelenmesi

28.11.2023

Sayın Burcu Özge RAZI

Yürütücülüğünüzü yapmış olduğunuz 2023-YÖNP-0815 nolu projeniz ile ilgili Bilimsel Araştırma ve Yayın Etiği Komisyonu'nun almış olduğu 23.11.2023 tarih ve 14/48 sayılı kararı aşağıdadır.

Bilgilerinize rica ederim.

KARAR 48- Sorumlu yürütücülüğünü **Prof. Dr. Aysun YAVUZ**'un yaptığı ve proje araştırmacısı **Burcu Özge RAZI** tarafından gerçekleştirilen "Ethical Implementation of AI in the Process of Academic Writing (Yapay Zekanın Akademik Yazım Sürecinde Etik Kullanımı)" başlıklı araştırmanın, Bilimsel Araştırma ve Yayın Etiği Komisyonu ilkelerine **uygun** olduğuna oy birliği ile karar verilmiştir.