KARADENİZ TECHNICAL UNIVERSITY * THE INSTITUTE OF SOCIAL SCIENCES

DEPARTMENT OF WESTERN LANGUAGES AND LITERATURE

APPLIED LINGUISTICS MASTER'S PROGRAM

COMPARING AUTOMATED AND TEACHER FEEDBACK: EFFECTS ON L2 WRITING ANXIETY, L2 WRITING MOTIVATION, AND STUDENTS' PERCEPTIONS

MASTER'S THESIS

Mehmet Akif YILDIRIM

JUNE - 2024

TRABZON

KARADENİZ TECHNICAL UNIVERSITY * THE INSTITUTE OF SOCIAL SCIENCES

DEPARTMENT OF WESTERN LANGUAGES AND LITERATURE

APPLIED LINGUISTICS MASTER'S PROGRAM

COMPARING AUTOMATED AND TEACHER FEEDBACK: EFFECTS ON L2 WRITING ANXIETY AND MOTIVATION, AND STUDENTS' PERCEPTIONS

MASTER'S THESIS

Mehmet Akif YILDIRIM

Supervisor: Asst. Prof. Dr. Hasan SAĞLAMEL

JUNE - 2024

TRABZON

APPROVAL

Upon the submission of the dissertation, **Mehmet Akif Yıldırım** has defended the study **"Comparing Automated and Teacher Feedback: Effects on L2 Writing Anxiety, L2 Writing Motivation, and Students' Perceptions"** in partial fulfillment of the requirements for the degree of Master of Arts in Applied Linguistics at Karadeniz Technical University, and the study has been found fully adequate in scope and quality as a thesis by **majority** on **19.7.2024.**

Committee Member		**Decision**		**Signature**
Name and Surname	**Mission**	**Accept**	**Refuse**	
Asst. Prof. Dr. Hasan SAĞLAMEL	Supervisor	☐	☐	
Prof. Dr. M. Naci KAYAOĞLU	Member	☐	☐	
Assoc. Prof. Dr. Şakire ERBAY ÇETİNKAYA	Member	☐	☐	

Approval of the Graduate School of Social Sciences.

Prof. Dr. Tülay İLHAN NAS
Director

DECLARATION OF ORIGINALITY

I, **Mehmet Akif YILDIRIM**, hereby confirm and certify that;

- I am the sole author of this work and I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution,
- this work contains no material that has been submitted or accepted for a degree or diploma in any other university or institution,
- all data and findings in the work have not been falsified or embellished,
- this is a true copy of the work approved by my advisor and thesis committee at Karadeniz Technical University, including final revisions required by them.
- I understand that my work may be electronically checked for plagiarism by the use of plagiarism detection software and stored on a third party's server for eventual future comparison,
- I take full responsibility in case of non-compliance with the above statements in respect of this work.

Signature……..

Date 2024

In loving memory of my grandfather, Ali Baş.
I wish you could see these days.

ACKNOWLEDGEMENT

I am honoured to be concluding this study with the memories of those who have provided me with immense motivation, inspiration and knowledge. Those who have contributed to the production of this thesis and the shaping of my academic development are highly appreciated.

I wish to express my profound gratitude to Asst. Professor Hasan Sağlamel, whose illuminating guidance, constant companionship, and friendly nature have been of immeasurable value throughout the recent chapters of my life. It is his genuine sincerity and dedication that have continuously inspired me to persevere, regardless of the obstacles that have come my way. Additionally, I have had the distinct honour of presenting my thesis to two other highly influential figures in my academic journey, my esteemed jury members, Professor Mustafa Naci Kayaoğlu and Associate Professor Şakire Erbay Çetinkaya. I feel deeply compelled to extend my heartfelt appreciation to both of them. Their significant contributions to my academic growth, both in the realm of knowledge and in terms of inspiration and motivation, have profoundly shaped my intellectual development and played a fundamental role in molding me into the academician I am today.

I wish to express my heartfelt gratitude to Dr. Lecturer Zeynep Öztürk Duman for her support and guidance throughout the course of this study. Additionally, I extend my sincere thanks to the students from the 2023 A and B preparatory classrooms, whose participation and contributions have profoundly enriched this journey.

June, 2024 Mehmet Akif YILDIRIM

TABLE OF CONTENTS

CHAPTER THREE

CHAPTER FOUR

ÖZET

Bu çalışma, Karadeniz Teknik Üniversitesi İngiliz Dili ve Edebiyatı Bölümü'nde okuyan hazırlık sınıfı öğrencilerinin bakış açılarını ortaya koymakta ve otomatik ve öğretmen geribildiriminin yabancı dil öğrencilerinin yazma kaygısı ve motivasyonu üzerindeki etkilerini araştırmaktadır. Bu amaçla, İkinci Dil Yazma Kaygısı Envanteri ve İkinci Dil Yazma Motivasyonu Ölçeği kullanılmıştır. Çalışmaya katılan 73 öğrenciden bir grup öğretmen geribildirimi alırken, diğer grup ChatGPT ve Paperpal'dan altı hafta boyunca paragraf yazma ödevlerine ilişkin yapay zeka tarafından üretilen geribildirim almıştır. Paperpal biçimle ilgili geribildirim verirken, ChatGPT içerikle ilgili yorum ve önerilerde bulunmuştur. Araştırmanın sonunda iki grup arasında bir son test yapılmış ve her gruptan altı katılımcıyla yapılan yarı yapılandırılmış görüşmelerde katılımcıların yazma geribildirimi yöntemiyle ilgili deneyimleri, kaygı ve motivasyon düzeyleri hakkında nitel bilgiler elde edilmiştir. Nicel bulgular, her iki geri bildirim yönteminin de öğrencilerin kaygı düzeylerini, özellikle de bilişsel kaygılarını önemli ölçüde azalttığını göstermiştir. Uygulama sonrasında, her iki grup da motivasyonda orta düzeyde bir düşüş yaşamış, içsel motivasyon önemli ölçüde azalırken dışsal motivasyon düşük ve sabit kalmıştır. Gruplar arasında içsel ve dışsal motivasyonda anlamlı bir fark bulunmamasına rağmen, deney grubu kontrol grubuna kıyasla motivasyonsuzlukta bir artış sergilemiştir. Nitel analiz, otomatik geribildirim grubunda, özellikle geribildirim detayı, çabukluk ve olumlu yorumlarla ilgili olarak, yazma isteğinin arttığını, kaçınmanın azaldığını ve yüksek memnuniyet olduğunu ortaya koymuştur. Buna karşılık, öğretmen geribildirim grubu, özellikle geribildirim detayı ve hızı konusunda daha düşük memnuniyet sergilemiştir. Çalışma, motivasyon ve kaygıya ilişkin öngörülerin yanı sıra, öğrencilerin yapay zeka araçlarını kullanımını ve bunlara yönelik tutumlarını da araştırmakta ve gelecekteki çalışmalar için öneriler sunmaktadır.

Anahtar Kelimeler: İkinci Dilde Yazma, Otomatik Geri Bildirim, Yapay Zekâ Geri Bildirimi, Öğretmen Geri Bildirimi, İkinci Dilde Yazma Motivasyonu, Yabancı Dilde Yazma Kaygısı

ABSTRACT

This study investigates the effects of automated and teacher feedback on EFL students' writing anxiety and motivation, while capturing the perspectives of preparatory year students studying at the Department of English Language and Literature at Karadeniz Technical University. To this end, the Second Language Writing Anxiety Inventory and the Second Language Writing Motivation Scale were used. Of 73 learners, one group received teacher feedback, while the other received AI-generated feedback from ChatGPT and Paperpal on their paragraph writing assignments over six weeks, with Paperpal providing feedback on form and ChatGPT offering comments and suggestions on content. After the experiment, a post-test compared the two groups, and semi-structured interviews with six participants from each group offered qualitative insights into their experiences with the writing feedback method, anxiety, and motivation levels. The quantitative findings indicated that both feedback methods significantly reduced students' anxiety levels, particularly cognitive anxiety. Post-intervention, both groups experienced a moderate decrease in motivation, with intrinsic motivation significantly declining and extrinsic motivation remaining low and stable. Although no significant differences were found in intrinsic and extrinsic motivation between the groups, the experimental group exhibited an increase in amotivation compared to the control group. The qualitative analysis revealed increased willingness to write, reduced avoidance, and high satisfaction with the automated feedback group, particularly regarding feedback detail, promptness, and positive comments. In contrast, the teacher feedback group exhibited lower satisfaction, especially concerning feedback detail and speed. Besides insights into motivation and anxiety, the study also explores learners' use of AI tools and their attitudes towards them, offering suggestions for future research.

Keywords: L2 Writing, Automated Feedback, AI Generated Feedback, Teacher Feedback, L2 Writing Motivation, Foreign Language Writing Anxiety

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AEE	: Automated Essay Evaluation
AES	: Automated Essay Scoring
AI	: Artificial Intelligence
AWCF	: Automated Written Corrective Feedback
AWE	: Automated Writing Evaluation
AWF	: Automated Writing Feedback
EFL	: English as a Foreign Language
ELT	: English Language Teaching
ESL	: English as a Second Language
GenAI	: Generative Artificial Intelligence
GPT	: Generative Pre-Trained Transformer
KTU	: Karadeniz Technical University
L2WMS	: Second Language Writing Motivation Scale
LLM	: Large Language Model
OPF	: Online Peer Feedback
SLWAI	: Second Language Writing Anxiety Inventory
SPSS	: Statistical Package for the Social Sciences

INTRODUCTION

The capacity of written language to reveal aspects of an individual's development in ways not feasible through other means has been recognised by researchers (Brown, 1998; Horning, 1987). The skill of writing has become increasingly important across a range of professions and disciplines, particularly in the digital age, where the ability to convey emotions, ideas, or information effectively through writing is valued (Graham, 2022). Consequently, the significance of writing instruction has grown as it equips learners with the ability to communicate their ideas convincingly, professionally, and effectively in personal and professional situations. Advanced writing skills are critical for critical thinking and learning, as they enable the organization and clarification of ideas and thoughts. Writing instruction through means such as feedback can help individuals at any stage of their lives to improve their writing abilities, and it can take various formats, including formal seminars, workshops, and presentations, serving learners in numerous ways throughout their academic and professional careers as well as for leisure activities.

Despite the acknowledged significance of writing skills, teaching writing remains a challenging endeavor (Negari, 2011). Various factors contribute to these difficulties, including anxiety (Cheng, 2002; Dewaele & Al-Saraj, 2015), technological adaptation (Warschauer & Grimes, 2008), cultural and linguistic diversity (Lee & VanPatten, 2003), classroom management (Farrell, 2013), and teacher qualifications (Daly & Miller, 1975; Daly, 1979; Kiuhara et al., 2009; Phuket & Othman, 2015). Reducing class sizes could potentially mitigate some of these challenges by enabling teachers to devote more individualized attention to each student, thereby decreasing their overall workload (Sarman, 2022). However, the feasibility of this modification may be hindered by financial restrictions. Teaching writing presents considerable challenges due to elements such as rapid advancements in technology (Li et al., 2019), linguistic diversity, and varying instructor qualifications (Kalay & Büyükkarcı, 2020). In the midst of these complexities, traditional feedback systems have limitations, while individualized feedback emerges as a critical aid in improving students' writing skills and resolving errors. The shift to emergency remote learning has made it more challenging to provide effective writing instruction and feedback (Balıkçıoğlu-Akkuş & Altay, 2023). As a result, in order to improve writing pedagogy and promote skill development, it is necessary to investigate modern methods such as automated feedback technologies.

Feedback, as defined by Hattie and Timperley (2007), is information supplied by an agent concerning specific aspects of a task's execution. Although several definitions of feedback exist, it can be simply defined as information regarding one's performance or comprehension provided by

various sources such as instructors, classmates, parents, or the self, and serves as a consequential reaction to that performance (Hattie & Timperley, 2007). Writing proficiency is crucial for English language learners in order to attain academic success and professional advancement, as well as numerous other advantages. In the context of student success, learners' development is observed to be evaluated commonly through their writing skills (Graham et al., 2014; Korkmaz, 2015), while writing is also viewed as one of the best predictors of academic success (Noeth & Kobrin, 2007).

CHAPTER ONE

1. FRAMEWORK OF THE STUDY

1.1. Writing Feedback

Writing feedback can be provided in various ways and by various agents, such as instructors, peers, or one's self (Birjandi & Hadidi Tamjid, 2012). A diverse array of feedback methods has been developed by educators and researchers to enhance students' writing skills. One such method is corrective feedback, which involves identifying and rectifying errors while emphasizing grammatical precision (Bitchener et al., 2005). In contrast, directive feedback guides students in rectifying identified mistakes (Sanders, 2005). Beyond these individual approaches, interactive feedback engages students through comments, messages, or verbal exchanges. Evaluative feedback takes a broader perspective, aligning writers with instructional objectives and evaluating drafts against an ideal standard. It also emphasizes sentence-level errors while providing guidance for future improvement. Formative feedback is more personalized, tailoring its approach to individual writers through inquiry-driven methods. By questioning students, this method aims to deepen reader comprehension and encourage substantial revisions in subsequent drafts (McGarrell & Verbeem, 2007). In contrast to formative feedback, summative assessments are conducted at predefined intervals to evaluate students' overall advancement, proficiency, and knowledge acquisition (Graham, 2018). Finally, peer feedback involves students providing evaluations on each other's work in various settings, such as the classroom, homework assignments, or digital platforms.

In addition to the practices and methods discussed above, the timing of feedback is also crucial. Intermittent feedback, characterised by the regular delivery of feedback at specific intervals in the writing process, allows students to reflect on their work and make necessary revisions, thus ensuring that students have time to reflect on their work and make the necessary revisions. Conversely, continuous feedback provides immediate guidance, often facilitated by technological tools or direct interaction with the teacher. This method offers students the opportunity to address errors promptly and refine their writing skills in real time. While continuous feedback excels at identifying and correcting immediate problems, intermittent feedback provides opportunities for more comprehensive assessments. Ultimately, the optimal feedback strategy depends on several factors, including student needs, the complexity of the writing task, and available resources, as Kellogg et al. (2010) note.

1.2. Challenges in Writing Instruction

One significant issue in writing instruction is the prevailing test-driven nature of modern curricula. İçbay (2005) highlights that the disregard for overlooked components of the curricula, such as writing and speaking skills, can render them uninteresting and dysfunctional for students who tend to prioritize the areas tested in examinations. Aydın's (2008) study revealed that the test-driven nature of tertiary-level schools in Turkey generates anxiety in students as they are not given enough opportunities to practice their writing skills. This lack of preparedness, a critical factor contributing to language anxiety, can persist into university-level education, where students may be ill-prepared to write in English. Zerey's (2013) examination of writing anxiety among pre-service EFL teachers in Turkey found that a majority of the students exhibited high or average levels of writing anxiety. Furthermore, multiple studies underscore the inadequacy of teacher qualifications and preparedness as feedback providers and writing educators, as reported by students. Issues related to motivation and preparedness are also reported affect teachers in the instruction and assessment of writing (Kalay & Büyükkarcı, 2020). In secondary education, reported difficulties in teaching English writing included students' poor grammatical competence, challenges in developing writing ideas, inadequate vocabulary knowledge, lack of motivation to learn, and teachers' poor time management, limited access to resources and facilities (Amalia et al., 2021).

Time is a highly valued resource in education, and feedback provision imposes several requirements on the providers. Instructors are often burdened by the time-consuming nature of assessment and feedback, especially when they need to provide feedback for multiple classrooms of students' essays, each of which may be thousands of words long (Ferris, 2007; Kearns, 2012; Liu et al., 2016; Han & Sarı, 2022). The incorporation of multiple drafts only exacerbates this issue for instructors. Providing optimal feedback involves not only highlighting errors and suggesting ways to correct them but also explaining the mistakes and offering guidance on how to avoid them in the future. Feedback that includes an explanation of a specific correction and provides recommendations for enhancing the text has been found to be more effective than feedback that only points out errors without further clarification (Shute, 2008; Liu, 2024). In addition, as scholars such as Graham and Perin (2007) and Liu (2024) report, giving feedback should not only involve pointing out mistakes, but also offering praise. To ensure students' optimal learning outcomes, feedback must be timely and provided promptly after submission. Delayed feedback can result in students losing sight of the subject matter, arguments, and ideas presented in their essays, which can diminish their engagement and motivation (Gibbs & Simpson, 2004; Fu & Li, 2022).

Among the issues experienced with EFL instruction, one of the most significant would be the limited exposure to the language (Alaraj, 2016; Amalie et al., 2021). This inadequate exposure can impede the growth of communicative proficiency and, in turn, have an impact on the learner's capacity to successfully use the language in their daily lives as well as academic and professional

settings (Al Hosni, 2014). Learners of English as a second language (ESL) enjoy the benefit of frequent exposure to the language they are learning through various modes such as speech, written text, and media. In contrast, learners of English as a foreign language (EFL) face limitations in terms of exposure to the target language, albeit with exceptions presented by technology (Wang et al., 2012; Chun et al., 2016; Yoestara & Putri, 2019). EFL learners' writing skills are no exception to this handicap, which could hinder the development of writing skills and lack of motivation on the parts of both teachers and students (Cheng, 2002).

1.3. Technology in Education

In recent years, technology has become an integral part of education with various advancements in both hardware and software technology. Tools such as computers, projectors, screens, and smart boards have become commonplace in classrooms, paving the way for several improvements in education. Alongside this hardware, websites and applications such as Google Classroom, Microsoft Word, Google Drive, and social media platforms including Facebook, Instagram, and TikTok are being tested for use in education (Hew & Cheung, 2014; Manca & Ranieri, 2016; Selwyn, 2016). Recent developments in software technology have allowed for the creation of software capable of learning through a set of data gathered from various sources. Machine learning and artificial intelligence, although in their early stages of development over the past decade, have proven to be instrumental tools for various branches of science and development. In education, artificial intelligence is being utilized by software developers to create apps and websites capable of using data provided by users to suggest what students need to learn. For instance, personalized learning programs such as Duolingo have utilized artificial intelligence to create a personalized learning path based on a user's strengths and weaknesses (Handini et al., 2022).

Regarding writing in English as a foreign language (EFL), the internet has been utilized by both students and teachers for decades. Various tools such as thesauruses, dictionaries, and translation and rephrasing tools have been widely utilized to improve writing skills. However, recent developments in software technology have paved the way for more sophisticated and personalized writing tools. For instance, tools such as Grammarly and ProWritingAid use artificial intelligence to analyze the writing of the user and provide suggestions for improving grammar, punctuation, and style. Moreover, the integration of technology in education has resulted in a shift towards a more student-centred learning approach (Murthy et al., 2017). For instance, online learning platforms such as Udemy and Coursera have enabled learners to access educational resources at their own pace and convenience. Additionally, collaborative online tools like Google Drive and Dropbox have facilitated group projects, making it easier for students to work together regardless of their geographical locations (Larson et al., 2017).

In conclusion, technology has played a crucial role in shaping modern education. The integration of sophisticated hardware and software tools in classrooms has facilitated personalized learning, improved writing skills, and enabled learners to access educational resources at their convenience. As technology continues to evolve, its impact on education is expected to increase, providing endless possibilities for both teachers and students alike.

1.3.1. Artificial Intelligence

Having gained increased popularity in the recent decade, artificial intelligence (AI henceforth) is displaying its potential in a variety of endeavours. A report by Chui et al. (2018) indicates that AI could potentially create between $3.5 and $5.8 trillion in value annually across nine business functions in 19 industries. In the field of medicine, AI is already being used to analyze patterns of symptoms and diagnose diseases efficiently and quickly. In agriculture, AI is helping to predict crop requirements by analysing variables such as soil conditions and weather patterns, leading to higher crop yields and reduced waste (Javaid et al., 2022). In physics, AI is being used to analyze data from large experiments and develop new methods and models through the analysis of complex data sets and simulations (Xu et al., 2021). In entertainment, AI is leading to the development of personalized content and smart computer games that respond to the player's actions and decisions. In education, AI is revolutionizing the way students learn and interact with information. AI technology is being used to analyze students' strengths and weaknesses, providing personalized learning experiences based on individual needs (Akyuz, 2020; Jaiswal & Arun, 2021; Maghsudi et al., 2021). Indeed, the instances mentioned above are only a small fraction of the whole range of possibilities for AI in various fields. Indeed, language education and linguistics were within the foci of AI developers and users as well. Several mobile and computer applications such as Duolingo, Busuu and ElsaSpeak were developed and published to provide users with personalized learning experiences, offer an experience that is supported by a set of courses that modify in difficulty and content based on the users' performance. In addition, automated writing evaluation tools utilize AI to offer students feedback on their writing assignments. They check the writing for quality of content as well as grammatical, spelling, and punctuation problems (Hockly, 2019).

The potential limitations and challenges of AI have been widely discussed in academic literature. One of the most significant issues is the lack of transparency in decision-making due to the complexity of AI models (Goodman & Flaxman, 2017). It is crucial to ensure that AI systems are trained on unbiased and adequate data, as biased or inadequate data could result in flawed predictions and judgments. The ethical and social issues associated with AI, such as privacy and security, have also been extensively explored. Moreover, there is a possibility of AI systems failing or making bad decisions, which could have significant negative consequences as they become more autonomous and complex. Finally, human oversight is still necessary to ensure the proper and efficient use of AI, even if it could automate various tasks (Snow, 2018, July 26). To overcome these

challenges and minimize the associated risks, various stakeholders, including policymakers, researchers, and industry professionals, must collaborate and engage in continuous research and innovation projects.

1.3.2. Artificial Intelligence in Education

Artificial intelligence holds the potential to revolutionize the field of education. AI is able to provide students with individualized learning experiences that are suited to their unique personality traits by employing machine learning algorithms. This process enables effective learning outcomes by facilitating increased engagement and the retention of knowledge. Likewise, the application of AI can make grading easier by delivering more accurate and objective evaluations of student work, reallocating teachers' time to concentrate on education. Additionally, AI tools can assist instructors in recognizing students who may be experiencing academic stress or a lack of motivation, enabling timely guidance and necessary adjustment. The benefits of AI in education are substantial, and as technology continues to advance, further innovative applications of AI to enhance the learning experience for students can be anticipated. The way individualized learning is approached could be completely changed by AI-powered education. AI has been integrated into an expanding variety of educational applications in recent years to deliver students individualized feedback, recommendations, and assessments (Duong et al., 2019; Upadhyay & Khandelwal, 2019; Luan et al., 2020). For instance, AI algorithms are used by language-learning apps such as ELSA Speak and Duolingo to adjust to the competence levels, interests, and learning preferences of users and offer real-time corrections and explanations (Handini et al., 2022). Moreover, web extensions such as Grammarly, which utilize AI to identify and remedy grammar mistakes in real time, can benefit learners by honing their writing skills and eliminating language obstacles (Fitria, 2021). LLMs such as ChatGPT, Gemini and Copilot have also been explored in their abilities to assist or provide in the provision of language education (e.g. Kasneci et al., 2023; Li et al., 2023; Zhou, 2023 Vashishth et al., 2024). The aforementioned and additional uses show that AI has the ability to improve education's efficacy and efficiency while also ensuring that it is easier to access and inclusive for students of various abilities and backgrounds.

Effective communication and cultural understanding have always been important aspects of language learning. In the contemporary digital age, technology has revolutionized the approach to language learning (Warschauer, 2002). Over the past years, language learning has witnessed a significant shift towards digital tools and resources such as online platforms, mobile applications, and gamification methods. These digital resources allow learners to participate in genuine language-related activities, communicate with native speakers, and receive immediate feedback on their language performance, thereby enhancing the learning experience (Stockwell, 2012). In the contemporary digital era, technology has brought about a paradigm shift in the approach to language learning (Chapelle, 2010; Dudeney & Hockly, 2012). Learners are now provided with a wide range

of digital tools and resources, such as mobile applications, online platforms, and gamification techniques, which offer opportunities to engage in authentic language tasks, communicate with native speakers, and receive real-time feedback on their language performance (e.g., Stockwell, 2012; Welbers et al., 2019). The theory of proximal development posits that learners can engage in activities that are just beyond their current level of ability, with the assistance of a more knowledgeable other, to facilitate learning (Vygotsky & Cole, 1978). In this regard, technology has the potential to offer learners an environment where they can engage in tasks that suit their individual learning needs and abilities. Empirical evidence has established the potential of technology to offer learners a personalized learning experience, leading to substantial enhancements in language learning outcomes (Polly & Byker, 2020). The promotion of autonomy in learners is equally important in language learning as they must develop the skills necessary for self-directed learning. According to Little (1991), learners who possess a high degree of autonomy are better equipped to take responsibility for their own learning, set realistic learning goals, and seek out resources that best suit their individual learning needs. Additionally, autonomous learners are more likely to engage in self-reflection and self-evaluation, leading to a deeper understanding of their own learning processes and improved language proficiency (Benson, 2013). In this sense, promoting learner autonomy can lead to more effective and efficient language learning outcomes. Task-based language learning and communicative language learning are also among the influential approaches in language education. Task-based language learning emphasizes the use of authentic tasks for learning and development (Ellis, 2003), while communicative language learning underscores the importance of using language for meaningful communication in real-life situations (Savignon, 1991). Through the delivery of tailored content, continuous practice and individualised feedback with the help of digital tools and AI assistants, the use of technology in language learning may align with these approaches, as digital tools and resources provide learners with opportunities to engage in authentic tasks and communicate meaningfully (e.g. Cao et al., 2024; Zhang, 2024).

Recent research has demonstrated that technology in language learning can enhance learners' motivation and engagement (Stockwell, 2010). Gamification, for instance, has been found to be an effective strategy for engaging learners in language learning activities. Digital tools such as videos, podcasts, and online communities have been found to foster a sense of community and social support, leading to increased motivation to learn (Kay, 2012; Schwarz & Caduri, 2016; Nacak et al., 2020). The shift to emergency remote education due to the COVID-19 pandemic has resulted in challenges for language learners, particularly in terms of motivation, anxiety, and lack of social interaction (Hazaea et al., 2021; Skar et al., 2021; Wilson & Lengeling, 2021). Language learning is a social process that involves communication with others, and the lack of face-to-face interaction in remote learning environments can negatively affect learners' motivation and engagement (Huang et al., 2020). In contemporary remote language learning, the dearth of social interaction and communication can potentially undermine learners' motivation and engagement. However, the integration of technological tools, such as online forums, chat rooms, and video conferencing,

showed promise in mitigating these issues. While some of these tools may not be as prevalent as they were in the past, they still hold significant value as means of promoting social interaction and communication among remote learners. Recently, emerging digital tools, such as virtual reality and social media platforms, have demonstrated efficacy in facilitating social engagement and collaboration in language learning (Khoshnevisan & Le, 2018; Koole et al., 2010). Consequently, the adoption of such tools can potentially enrich the remote language learning experience and foster social interactions crucial to the language learning process. Additionally, the use of game-based learning and interactive digital materials can increase learners' motivation and engagement, providing a more enjoyable learning experience (Plass et al., 2015 Peterson, 2016). Therefore, the integration of technology into language learning can be a valuable resource for both learners and instructors in the transition to remote learning environments.

The integration of AI has fundamentally transformed language learning in the digital age. AI technologies, such as natural language processing and machine learning algorithms, have been designed to provide personalized language learning experiences (Yang et al., 2021). These technologies enable learners to receive immediate feedback on their language performance and customized language learning materials based on individual learning needs.

1.3.3. Language Learning with AI

In recent years, there has been a rise in the development of innovative AI-powered language learning platforms such as Lingvist®, Duolingo®, and Busuu®. Lingvist, for instance, employs AI algorithms to provide learners with a personalized language learning experience, including adaptive exercises and real-time feedback on their language performance. This platform offers an intelligent learning system that can assess learners' needs and provide them with tailored learning materials, which can enhance their language skills. Duolingo®, one of the most popular language learning apps, leverages AI algorithms to track learners' progress and provide personalized feedback based on their performance data. The app uses a combination of natural language processing and machine learning to adapt to learners' strengths and weaknesses and provide tailored practice exercises to reinforce learning (Hagiwara, 2021). The system also provides a community of learners, allowing them to interact and practice their language skills with others in a supportive environment. These innovative language-learning platforms have been proven to be effective in providing learners with a personalized and engaging language-learning experience, resulting in significant improvements in their language-learning outcomes (Gangaiamaran & Pasupathi, 2017; Godwin-Jones, 2011).

Recent advances in technology have led to the development of AI-powered language learning tools, including chatbots and language translation tools. Chatbots are designed to provide learners with conversational practice in a simulated real-life context (Smutny & Schreiberova, 2020). Moreover, AI-powered language translation tools, such as Google Translate, have become

increasingly sophisticated in recent years, providing learners with instant translation services for a wide range of languages (Baydere, 2023). In recent years, the popularity of AI-powered conversational agents designed to assist language learners in their practice has increased significantly. One notable example is ChatGPT, developed by OpenAI, which employs advanced language models and natural language processing techniques to provide personalized language practice and feedback to learners (OpenAI, 2021). These tools have the potential to revolutionize language learning by offering learners personalized and interactive language practice experiences that simulate real-life language use. AI technologies can offer learners real-time feedback, personalized learning materials, and immersive language practice opportunities (e.g. Cao et al., 2024; Yaşar, 2020). As the capabilities of AI continue to advance, and as the possibilities for language learning are limitless, it is crucial for educators and learners to exploit such innovative technologies.

1.3.4. Automated Writing Evaluation

The integration of technology in the field of language education is not a new phenomenon, as it has been explored and implemented in various ways for decades (Warschauer & Healey, 1998). However, with the rapid advancement of AI technology, innovative tools and solutions for language education have become increasingly prevalent in recent years (Liu & Ren, 2022). Automated writing evaluation (AWE) is a type of AI technology that allows for the analysis and feedback of written work, such as essays, reports, and articles, through the use of natural language processing (NLP) techniques. These techniques involve the examination of various linguistic aspects, including syntax, grammar, vocabulary, and discourse, to identify errors and provide suggestions for improvement. AWE tools may also incorporate machine learning algorithms to adapt to the writing styles and preferences of individual users and offer personalized feedback. In essence, AWE technology provides a means to enhance the writing skills of language learners by offering immediate and personalized feedback on their writing. This can be particularly beneficial for English as a second language (ESL) or English as a foreign language (EFL) learners who may face challenges in developing their writing skills due to limited exposure to the target language. Through the use of AWE tools, learners can receive targeted feedback on specific areas of improvement, leading to enhanced writing proficiency over time. A variety of AWE tools are present in the current date, with examples such as E-rater, Turnitin, Criterion, Grammarly, and WriteLab.

While there are a multitude of benefits of AWE tools, they are not without flaws. Compared to human readers, automated writing evaluation methods possess multiple critical limitations. Even though AWE tools very quickly and accurately evaluate written work in terms of grammar, spelling, grammar, vocabulary and discourse, they currently lack in terms of content evaluation (Crossley et al., 2014). Language features such as puns and irony may be overlooked by AI-based evaluators as well (Weigle, 2013). To name a few, the existence of constructive, though inadequate comments and suggestions, which were perceived to be unclear, were criticized in the study of Li et al. (2023). On

the other hand, Deane (2013) criticized the lack of feedback on argumentation or rhetoric strength of texts, while Abubakir (2021) highlighted the effects of students' technological skills on te effectiveness of AWE use.

1.4. Definition of Key Terms

To ensure clarity and understanding, it is essential to define the terms used in this study. These terms are defined within the context of the study as follows:

Feedback: refers to the "information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one's performance or understanding" (Hattie & Timperley, 2007: 81)

Automated Feedback: refers to writing feedback provided by a computer software or AI tool (Shermis & Burstein, 2002). In this study, the feedback tools used were Paperpal and ChatGPT.

Teacher Feedback: within the confines of the present study, teacher feedback stands for any type of comment, suggestion or praise provided regarding the work of a student.

Motivation Scale: refers to the Second LanguageWriting Motivation Scale (L2WMS)(Eryılmaz & Yeşilyurt, 2022) that will be used for the purpose of the study.

Anxiety Scale: refers to the second language writing anxiety scale (SLWAI) developed by Cheng (2004), which will be used in this study to measure the participants' anxiety levels through pre- and post-test surveys.

AWE: refers to any kind of feedback or analysis provided by computerised or AI-based tools.

ChatGPT – refers to the AI-powered chatbot published by OpenAI which is able to analyse and provide computer-generated responses to user-generated text (https://chatgpt.com/).

Paperpal – refers to the web- and software-based automated feedback tool that provides grammar, Automated Writing Evaluation (AWE) utilizes artificial intelligence (AI) to assess and provide feedback on various types of written compositions, ranging from essays to reports. The tool encompasses functions that facilitate error correction in spelling and grammar (https://paperpal.com/).

1.5. Significance and Purpose of the Study

As the 2019 COVID-19 epidemic has accelerated the move toward remote learning and furthered the profound changes driven by the digital age in the field of education, there has been increased concern surrounding the management of students' motivation and anxiety levels. Instructors' ability to provide timely and accurate feedback is crucial for improving students' motivation and lowering anxiety levels (Fu & Li, 2022). The present research aims to investigate the performance of automated feedback in contrast to instructor feedback, evaluating its effects on students' motivation and anxiety levels in the context of writing instruction.

The significance of this research lies in its potential to provide insight into the ways automated feedback could increase learners' motivation and decrease their anxiety, which could have major implications for writing instruction, benefiting learners by offering insights on how writing instruction may be gauged for increased writing motivation and alleviated writing anxiety. This study further intends to contribute to the implementation of automated feedback systems that can help teachers deliver individualized and timely feedback to learners. Furthermore, feedback provision is a tedious and time-consuming burden on instructors to provide learners with timely and practical feedback. This study endeavours to address a research gap by scrutinizing and comparing the effects of automated feedback and teacher feedback on student motivation and anxiety levels in the context of English language instruction. In addition, this study also aims to shed light on the potential positive uses of the tool ChatGPT, as well as its limitations and associated risks. Likewise, the study seeks to explore the potential of Paperpal for providing writing feedback. The results will provide practical insights for language teachers who wish to benefit from automated feedback opportunities. With many tools and applications around, language teachers may not feel at ease to make well-informed decisions when giving feedback.

This study aims to investigate the impact of two feedback approaches: teacher-provided and automated (using Paperpal® and the large language model ChatGPT) on students' English writing anxiety and motivation. These aspects, although mentioned, require more detailed examination within the study. The significance of writing skills and the mastery of English grammar underscores the increasing importance of effective writing education. Typically, writing education relies on assignments such as paragraphs and essays, coupled with regular feedback on students’ errors and shortcomings. However, as class sizes grow, providing adequate feedback becomes a challenge for teachers. Hence, practical solutions are essential to enhance the efficiency of feedback methods. Notably, there's a dearth of research in Turkey comparing feedback methods' suitability, specifically between automated tools and teacher feedback. Therefore, this study aims to contribute to the literature by examining and comparing these two feedback methods concerning their impact on writing anxiety and motivation among students. By exploring the effectiveness of teacher feedback and automated tools like Paperpal and ChatGPT, this research intends to identify their respective

strengths and weaknesses. Additionally, it seeks to provide recommendations for refining feedback methodologies in English writing education.

Research Questions:

1. How do Automated Writing Evaluation (AWE) tools compare to teacher feedback in influencing students' motivation and anxiety levels in writing tasks?
2. Is there a statistically significant difference between the L2 writing motivation scores of students who received automated feedback and those who received teacher feedback?
3. Is there a statistically significant difference between the L2 writing anxiety scores of students who received automated feedback and those who received teacher feedback?
4. What are the perceptions of students towards the effectiveness of the two feedback modalities?

CHAPTER TWO

2. LITERATURE REVIEW

In recent years, the domain of education has undergone a significant paradigm shift as a result of the integration of technology, bringing in innovative tools and approaches that have demonstrated a discernible influence on pedagogical practices and educational outcomes. A significant and ongoing area of inquiry in the field of language education examines the relationship between NLP-enabled Artificial Intelligence (AI) technology and LLMs like ChatGPT, Gemini and Copilot with an emphasis on Automated Writing Evaluation (AWE) tools. This convergence not only shows how these technologies could improve students' writing performance, but also looks at how well they function to raise student engagement levels, lower anxiety, and encourage motivation. This research review aims to clarify the complex relationships between AWE, LLMs, language acquisition, student motivation, anxiety, and writing proficiency within the framework that has been provided. This assessment identifies gaps in the body of literature by a thorough examination of existing academic research and incorporates ideas from previous studies. The review also thoroughly synthesizes empirical data from several studies, shedding light on the complex processes behind AWE and LLM technologies. The review aims to promote an extensive discourse while acknowledging criticisms, vulnerabilities, and potential risks by carefully examining the numerous aspects of AWE and LLMs' usefulness, including their capacity to improve writing proficiency, anxiety, and motivation. As a result, this review aims to make a meaningful contribution to the continuing conversation on educational technologies while also providing an in-depth comprehension of the crucial functions that these tools have taken on within the context of students' academic experiences. The following discussion commences by introducing and providing an overview of both AWE and LLM technologies. Consequently, a detailed examination of scholarly arguments in favour of and against the utilization of these technologies ensues with a specific focus on prevalent academic perspectives and discussions. Subsequently, the review shifts to the subjects of motivation and anxiety, both of which have been observed to attract significant scholarly interest. Their examination within the context of AWE and LLMs is conducted by analysing pertinent studies. Finally, the discussion concludes by highlighting areas of research that require further exploration and summarizing the implications deduced from the collective body of literature.

2.1. Background and Context

Driven by continuous advancements in hardware and software technologies, the field of artificial intelligence (AI) is in a constant state of rapid evolution. The development of AI was forever altered by pivotal moments such as Alan Turing's creation of the Turing Test and John McCarthy's introduction of the term “artificial intelligence” during the groundbreaking Dartmouth Conference (McCarthy et al., 2006). The technology offers utilities in banking (Fares et al., 2022), healthcare (Murray et al., 2020), education (Beck et al., 1996; Holmes et al., 2023), and entertainment (Maes, 1995; Laird, 2002), and has the potential to revolutionize such and many other industries with the aid of continuous advancements and concurrent breakthroughs in computer and software technology. AI empowers computer systems to detect, identify and catalogue objects or patterns for various purposes (Miller & Brown, 2018), facilitate the training of advanced robots (Arulkumaran et al., 2017), and execute intricate tasks with heightened efficiency and precision. The popularity of AI technology and its potentials were demonstrated to the public in 1997 when Garry Kasparov, a renowned chess grandmaster faced defeat against computer program “Deep Blue”. This significant event showcased the profound impact and future potential of AI, showcasing its ability to surpass human capabilities and pave the way for groundbreaking advancements in various fields. As AI continues to revolutionize science and technology, it provides innovative solutions to issues that were previously thought to be insurmountable. The AI-powered assistants from science fiction films, such as HAL 9000 from “2001: A Space Odyssey” and Jarvis from the fictional Marvel superhero Iron Man are now being compared to real-world AI assistants such as Siri, Gemini, Amazon Alexa, and ChatGPT.

The review of the literature that follows delves into the constantly evolving setting of AWE along with AI tools and LLMs in the context of EFL writing feedback. Following an introduction to foundational research, the review delves into the advantages and applications of AWE and LLMs in various settings, focusing on their benefits in education and writing instruction. It also investigates ways for prompt engineering and discusses potential difficulties with the implementation of ChatGPT. The ethical considerations of AWE and LLMs are examined, as well as potential cases of created content misalignment. The review evaluates the dnaymics between these technologies and student motivation, anxiety, and foreign language anxiety in writing. It also assesses the efficiency of automatic writing feedback and correction tools.

2.2. ChatGPT

ChatGPT (https://chatgpt.com/) was released in 2022 by OpenAI, with newer versions being published and made accessible to the public frequently. On the 14th of March, 2023, the latest version of ChatGPT, GPT-4, was published. Newer abilities of GPT-4 included the ability to comprehend visual input, increased creativity and collaboration, and increased capability for context. While the

majority of the development of OpenAI's ChatGPT occurred in the past half-decade, there are numerous studies investigating its potential and effects (e.g.; Haleem et al., 2022).

2.2.1. The Operational Framework of ChatGPT

Natural language processing (NLP), a multifaceted collection of algorithms, software, and models designed to understand and generate human language in written form, empowers LLMs to unravel the complexities of human communication and bridge the gap between language and code. This field enables the translation of human language into machine-readable code and vice versa, facilitating a balanced exchange of information between humans and computers. The transformer architecture, which forms a fundamental component of ChatGPT's AI-based NLP, originated from the pioneering work of Google in 2017 (Vaswani et al., 2017). Google adopted this architecture to revolutionize their Google Translate tool, transitioning from a phrase-based machine translation system to a neural machine translation (NMT) system, resulting in significantly improved translation accuracy. OpenAI subsequently leveraged the same transformer architectures, as evidenced in the research conducted by Tay et al. (2022), and Vaswani et al. (2017), to develop ChatGPT (Vaswani et al., 2017; Tay et al., 2022). Natural language processing (NLP) was regarded as a key consideration in the assessment of writing with software, as evidenced by several studies (Kaplan et al., 1998; Grimes & Warschauer, 2010). In addition to NLP, alternative methods have been proposed for software to evaluate writing content and organization, including rule-based systems (Kaplan et al., 1998) and deep learning models (Crawford et al., 2023). These methods offer additional capabilities beyond NLP's reach, enriching the software's analytical toolset.

2.2.2. Prior Research on ChatGPT

Due to its many implications, uses and associated risks, ChatGPT has appeared within numerous studies, mostly as the subject, but rarely also as an author in the past half decade (e.g., Köbis & Mossink, 2021; Graham, 2022; Klekovkina & Denié-Higney, 2022; Zhai & Ma, 2022; Aktay et al., 2023; Biswas, 2023; Buriak et al., 2023; Choi et al., 2023; Crawford et al., 2023; Guo et al., 2023; Kasneci et al., 2023; O'Connor & ChatGPT, 2023; Rudolph et al., 2023; Wang et al., 2023; Yan, 2023; Zhou et al., 2023).

The research conducted by Hostetter et al. (2023) explored the attitudes and perceptions of 83 undergraduate students and 82 faculty members in the field of psychology. Participants were tasked with evaluating text generated by GPT-3 as well as text produced by undergraduate students. The majority of participants in the study showed limited ability to distinguish between AI- and human-generated content upon examination. Therefore, the overall frequency of detecting AI-generated content was found to be low. However, the study revealed that prior experience with GPT-3 played a significant role in enhancing participants' capacity to identify AI-generated content. These findings

shed light on the challenges associated with detecting AI-generated text and underscore the importance of familiarity and experience in accurately discerning between AI-generated and human-generated content. The findings also emphasize the importance of providing teachers with comprehensive knowledge and expertise in ChatGPT and similar tools. This proactive approach helps teachers detect plagiarism and cultivate a learning environment that fosters originality, academic integrity, and responsible use of AI-driven technologies. By embracing ChatGPT's potential and equipping educators with the necessary skills, a future where technology and human guidance coexist harmoniously fosters innovation, critical thinking, and ethical engagement in education.

In a recent qualitative study on the views of Turkish 4th-grade primary school students about ChatGPT by Aktay et al. (2023), the researchers observed that ChatGPT proved to be a useful source of accurate, timely, and comprehensible responses to students' inquiries throughout the experimental phase. Additionally, the students who took part in the study not only stated they enjoyed utilizing ChatGPT in their scientific classrooms, but they also showed favourable reactions to its introduction, which improved their attitudes toward learning. The findings derived from these studies, coupled with the reported advantages of ChatGPT, underscore the tool's promising potential as a valuable aid in education for both students and teachers. It has the capacity to alleviate anxiety levels among students and enhance their motivation, increase efficiency while studying for assignments, exams and encourage extracurricular activities. Additionally, it can promote student autonomy while simultaneously easing the workload burden on educators. The observed benefits of ChatGPT in educational settings suggest that it holds the potential to serve as a beneficial assistant, augmenting the learning experience and fostering a conducive environment for academic growth. Alafnan et al. (2023) conducted an empirical study on the potential and limitations of ChatGPT as an educational tool. They used both theory-based and application-based methodologies to explore its transformative capabilities in search engines. The study suggests ChatGPT has the potential to replace or revolutionize existing search engines, with advanced data analysis and user-generated relevant responses. This study emphasizes the impact of ChatGPT on improving search experiences and calls for further exploration of AI-driven technologies in information retrieval. While AI and NLP technology are already in use by major tech companies such as Google and Microsoft, language-model assisted information retrieval is being implemented into popular search engines such as Google and Bing.

These insights highlight the potential of ChatGPT as a valuable tool in language learning and open up avenues for further exploration and refinement to enhance its language recognition capabilities. ChatGPT offers valuable assistance to teachers in various aspects of their work. It is capable of generating lesson plans and formulating questions based on text and data. Additionally, ChatGPT can create frameworks and templates for different types of text formats, including application forms, resignation letters, and electronic mail. Other AI-based language processors and

models, such as Google's Gemini, Microsoft's Copilot, Meta's LLaMA, and Anthropic's Claude exist alongside ChatGPT, offering a range of options for various language processing tasks. The utilization of language processors has also shown potential in reducing foreign language anxiety and increasing willingness to communicate (Bao, 2019). Moreover, studies have been conducted to explore teachers' AI literacy and their readiness to embrace such systems (Choi et al., 2023; Chocarro et al., 2023). These investigations contribute to the understanding of the perceptions and attitudes of educators towards AI in the educational setting. Despite the remarkable advancements in AI, there are crucial aspects of human cognition that AI is unable to replicate or replace, as evidenced by previous research (Zhai & Ma, 2023). Recognizing that AI can be utilized in many areas of life is essential (Göçen & Aydemir, 2020; Buriak et al., 2023). However, there have been concerns raised on the reliability and accuracy of the data generated by LLMs. Buriak et al. (2023) draw attention to potential issues associated with GPT-generated content, including errors, fabricated sources and references, and the inclusion of irrelevant claims. Their findings underscore the need for a balanced perspective on the use of AI in education and research. While LLMs offer valuable support and efficiency, ethical considerations, the preservation of critical thinking abilities, and critical evaluation of AI-generated content must be prioritized. Continued research and thoughtful integration of LLMs into educational practices can ensure its responsible and effective implementation. The study conducted by Wang et al. (2023) examined the utilization of ChatGPT as an assistant for developing boolean queries in systematic and literature reviews. Through their experimentation and prompt development process, the researchers found that ChatGPT showed potential in assisting with the creation of Boolean queries. However, it was observed that ChatGPT generated varying results with the same prompts, impacting its consistency. While ChatGPT demonstrated promise in generating Boolean queries, there is a limited number of studies on its performance in comparison to state-of-the-art methods designed specifically for this purpose. These findings, along with similar studies highlighting the extensive knowledge and utility of ChatGPT, suggest that although the tool may not perform at a specific task as efficiently as well as dedicated tools, its potential can be significantly enhanced through careful preparation and integration. The results of this study may also indicate that, despite the benefits of LLMs in terms of speed and reliability, there are certain tasks for which human performance is superior to that of ChatGPT's. Despite its effectiveness, ChatGPT could fall short of human performance in some situations that call for complicated judgment, contextual interpretation, and sophisticated decision-making. Due to these drawbacks, it is critical to acknowledge the special cognitive capacities possessed by humans and to take into account the complimentary roles that AI technology can play in particular tasks. Despite ChatGPT's outstanding capabilities, it is important to recognize its limitations and take advantage of its advantages with human experience to receive optimal results. The study conducted by Zhou et al. (2023), where a group of forty Chinese students majoring in English participated, compared the narrative writing performance of ChatGPT with that of human writers. The researchers analysed and compared 500-word narrative texts generated by the students and ChatGPT using the coh-metrix text analysis tool. The findings indicate that ChatGPT outperformed human writers in terms of narrativity, word concreteness, and referential cohesion.

However, the students demonstrated superior performance to ChatGPT in terms of syntactic simplicity and deep cohesion. Results of the study align with the findings reported by Wang et al. (2023), illustrating the differential strengths and limitations of ChatGPT in the domain of narrative writing. While ChatGPT demonstrates proficiency in certain aspects of narrative generation, human writers exhibit superior abilities in terms of syntactic complexity and cohesive writing. Consequently, these findings underscore the need for caution and restraint when relying on LLMs, chatbot systems and AI technologies. Excessive dependence on such systems, without incorporating human insight, is deemed ill-advised (Fui-Hoon Nah et al., 2023). By acknowledging the distinct competencies and shortcomings of AI systems and human writers, a balanced approach can be fostered, harnessing the unique capabilities of each to augment the quality and efficacy of written communication. In a qualitative study conducted by Yan (2023), the impact of ChatGPT on the writing practices of Chinese L2 undergraduate students was examined. Through the analysis of eight participants' experiences and perceptions, which were captured through computer screen recordings, observations, and in-depth interviews, Yan explored the students' views on ChatGPT's automated text generation features. The findings revealed that while students acknowledged the advantages of ChatGPT's capabilities, they also expressed concerns about potential issues such as inequity and plagiarism.

2.2.3. Advantages and Applications of ChatGPT

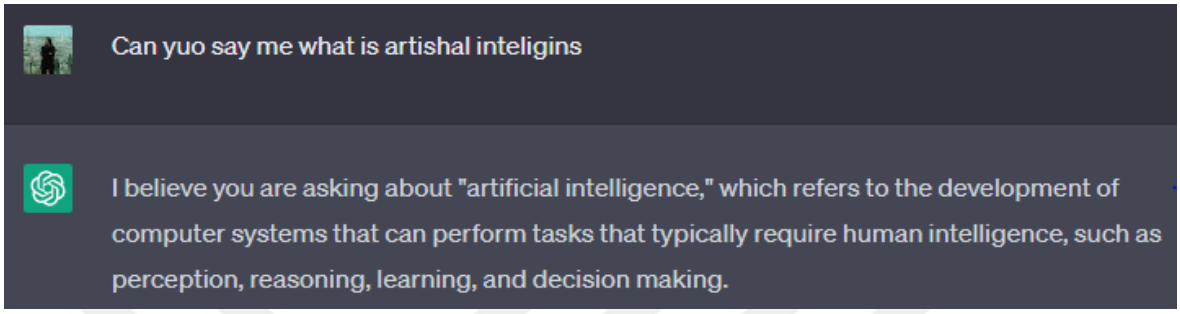
A thorough evaluation of the implications, risks, and consequences of incorporating recent technologies into education necessitates adopting a comprehensive perspective. This evaluation must consider not only the apparent potential benefits but also the associated risks and long-term feasibility. It should be conducted with a focus on methodological education, teacher training, and the educational environment and facilities (Cornu, 1995). Following a thorough examination of the issues, dangers, and safety precautions related to ChatGPT as documented in the literature, it is imperative to explore the potential benefits offered by this groundbreaking technology. Acknowledging and addressing the challenges and limitations of ChatGPT is essential for mitigating risks; however, it is equally critical to identify and evaluate the potential advantages users can gain from this technology. This section aims to elucidate the various positive applications of ChatGPT, thereby unveiling new opportunities for enhanced productivity, creativity, and accessibility. This section seeks to illuminate the various ways that ChatGPT can be used for good, opening up fresh opportunities for improved productivity, creativity, and accessibility. This section will focus on emphasizing the potential role of ChatGPT in determining the course of artificial intelligence and human-computer interaction by examining its potential applications across diverse domains. Furthermore, despite efforts to guarantee fairness, the limitations of EAS tools may indicate that they should be used in conjunction with human raters rather than as a substitute for them.

Biswas (2023) stated that by using chatbots and NLP, medical writers can potentially create more accurate and consistent documents faster and with less effort. The proposal of Crawford et al. (2023) states that ChatGPT could successfully be used for the purpose of creating essay plans, grammar checking, and receiving suggestions regarding missing information or content curation within the text, essentially serving students as a writing assistant. They also suggest that the use of ChatGPT could improve student autonomy, competence and relatedness. ChatGPT was observed to demonstrate effective applicability across diverse educational levels, including elementary schools, group and remote learning settings, and professional training (Kasneci et al., 2023). The researchers also suggest that LLMs can be modified to achieve certain goals and to perform in the most optimized form, but highlight the need for teachers and learners to develop competencies and literacies to comprehend the limitations. Ultimately, they conclude that LLMs hold promise in the transformation of language education and that their integration into education requires a cautious approach. ChatGPT and similar LLMs have the capability to assist students in organizing their writing and comprehending lengthy materials by generating summaries and outlines. Although ChatGPT lacks the critical thinking and reading abilities of human raters, its advantage lies in its access to extensive datasets, allowing it to excel in analyzing and producing text within specific domains such as chemistry, engineering, and medicine (Biswas, 2023). While ChatGPT may not possess the medical expertise of a qualified doctor, or that of an engineer, it surpasses human capabilities in processing and summarizing articles on various topics and branches of science all at once, such as nuclear physics, law and astrophysics. This highlights the notion that in a room of engineers, ChatGPT would function as the most proficient medical doctor, and vice versa in a room of medical doctors, where it would excel as the finest engineer.

While Hong (2023) suggests that ChatGPT should be examined in terms of student motivation and how it impacts the affective filter of students, he draws attention to the phase of adaptation learners and teachers are going through with this technology and concludes that once the initial flurry of avoidance has subsided, teachers will begin to notice the advantages and opportunities that ChatGPT offers. Haleem et al. (2022) explore ChatGPT's many functions and advantages, highlighting its capabilities and capacity to produce suggestions on a diverse set of content categories. The AI is praised for its ability to correctly distinguish AI-generated content after being trained on human-authored texts. ChatGPT's flexibility is further demonstrated by the way in which it may be used for a variety of linguistic inputs and scenarios, such as customer service, education, and entertainment. Future conversational AI systems stand to benefit from the model's capacity to learn and advance over time. The overall transformational role of ChatGPT in research, education, and diverse tasks emphasizes its potential significance. The inclusion of extraneous elements that would be refered to as “noise” such as grammar errors, typos, punctuation mistakes, and misused words in student essays, particularly among developing English learners, were concerned to present a challenge for AI systems to comprehend, evaluate, and provide feedback for (Grimes & Warschauer, 2010). However, through the implementation of NLP technologies and databeses used

by software and websites such as Grammarly and Quillbot, AI tools may well be able to successfully comprehend and provide feedback for student essays and text written by developing English learners. For instance, ChatGPT in its current state with the publicly available version is able to understand texts with spelling and grammar errors.

Figure 1: ChatGPT's Ability to Comprehend Text With Errors



The previous criticisms of AI on its inability to understand text with such noise may have become obsolete in that sense.

The studies examining the role and impact of ChatGPT have provided insights into its potential applications and implications. Talan and Kalınkara (2023) conducted a study comparing ChatGPT's performance with undergraduate students in an anatomy course. The outcomes showed that ChatGPT performed better than the students, highlighting its potential as a teaching tool. The study did, however, illustrate the AI's limits with regard to visual cue interpretation as well as the necessity of asking precise questions in order to receive appropriate answers. Similarly, Aktay et al. (2023) examined students' perspectives on integrating ChatGPT into their learning experience. The qualitative study with fourth-grade students revealed a positive view of ChatGPT, enhancing engagement, academic achievement, and information accessibility. Students recommended its use in various subjects, recognizing its potential to improve learning experiences. On the other hand of the discussion, Fırat (2023) investigated the integration of AI technologies into Learning Management Systems (LMS) for improved e-learning, with ChatGPT being one among these technologies. The study offered step-by-step examples and practical applications, emphasizing the significance of careful planning and evaluation when incorporating AI into LMS, considering factors such as privacy, ethics, pedagogy, and technical aspects. Ensuring data security, adhering to ethical principles, and aligning with educational goals were identified as vital components for successful AI integration in LMS settings. Moreover, Topsakal and Topsakal (2022) proposed a framework for developing foreign language teaching software for preschoolers, incorporating AR, Voicebots, and ChatGPT. The study highlighted the potential positive impact of early foreign language learning, fostering better communication and understanding among children. The research emphasized the effectiveness of the software in creating an engaging and entertaining learning environment, with

ChatGPT efficiently preparing interactive dialogues hosted in the software. The scalability of the software offered various language options and potential benefits to millions of children. Students who prefer a hands-on, personal learning experience will benefit from the use of language models and computer assessment and feedback tools. Students could also use chatgpt and similar language models for a plethora of good purposes such as to study for exams, discover new tools, practices and tools to improve their daily and academic lives (Alafnan et al., 2023).

The utilities of AI technology and LLMs in education extend beyond language education, encompassing a plethora of other fields that offer diverse and substantial benefits. Notably, fields including medicine (Eysenbach, 2023; Talan & Kalınkara, 2023; Temsah et al., 2023), software development and web design (W. Ma, et al., 2023) use AI and LLMs to achieve benefits that could possibly outweigh those observed in language education and learning. In medical education, Wood et al. (2021) explore the integration of AI literacy into curricula, advocating for the education of both students and faculty members regarding the benefits of AI. This concept of ChatGPT as a writing assistant for medical writing is further corroborated by findings from other researchers. (e.g. Biswas, 2023; Temsah et al., 2023). Similar findings are presented in the study of Ma W. et al. (2023) where ChatGPT is examined in terms of its ability to comprehend codes in various coding languages such as C#, Java, and Python, reporting the presence of hallucinations, which are instances where the LLM may present irrelevant or incorrect information, during the interpretation of codes. These applications include code generation and debugging, automated software testing, and writing assistance. The study also highlights certain limitations and flaws requiring thorough investigation and fine tuning. Overall, LLMs are praised in their ability to aid computer and software engineers in automating challenging and time-consuming tasks, improving algorithms, boosting data analysis, and accelerating the creation of web interface code.

Along with coding and engineering, ChatGPT was explored in its ability to assist in writing tasks as well. Dönmez et al. (2023) explored the challenges and opportunities of using ChatGPT as a co-writer in academic research. The research investigated whether chatbots could assist researchers in the scientific research process. Although creative ideas and numerous points of view were highlighted as advantages of AI technologies, issues with content authenticity and ethical ramifications—specifically, plagiarism—were highlighted. The study concluded that AI could provide support in tasks like title and summary generation but emphasized the need for further research to assess the capabilities and limitations of AI-driven software in academic publishing. Baloğlu and Çakalı (2023) focused on the implications of incorporating AI tools, specifically ChatGPT, in academic research and its potential effects on academic ethics. The study used ChatGPT to generate textual content, showcasing its capabilities in producing academic materials that may bypass plagiarism detection measures. The research emphasized the importance of examining potential violations of academic ethics in AI-generated content, particularly regarding issues such as proper citation and originality. The study urged the development of technologies to identify AI-

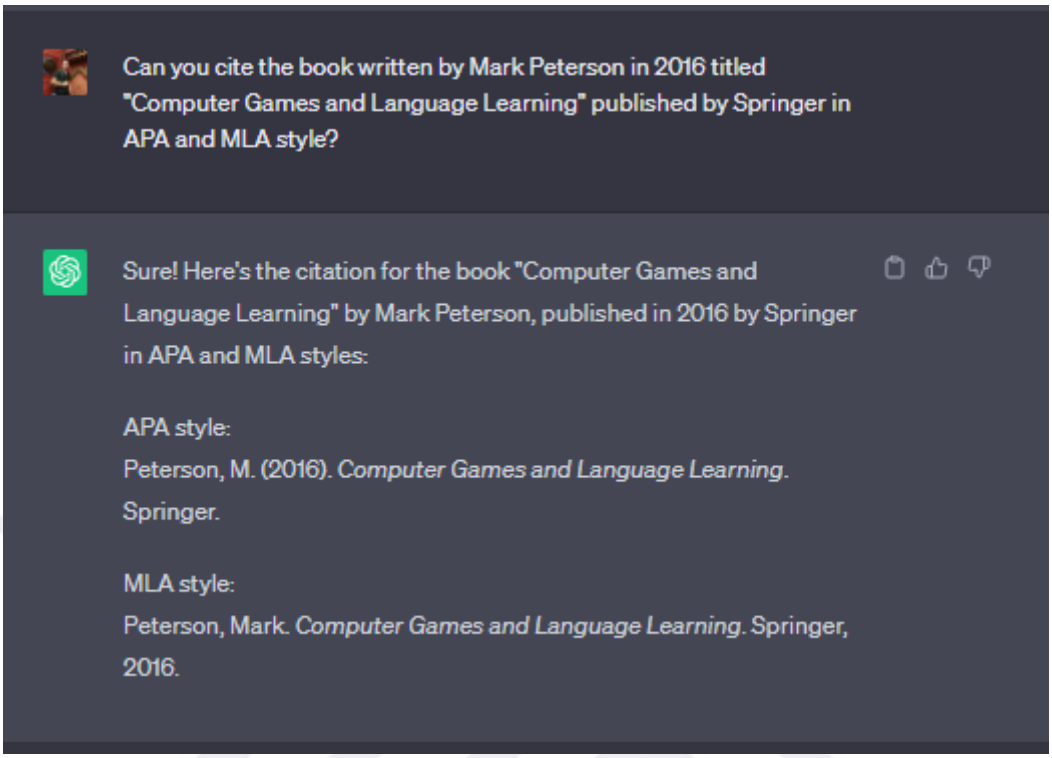
generated work and advocated for educational initiatives to reinforce ethical standards among researchers, promoting responsible and transparent practices in academic writing. The study of Aljanabi et al. (2023), co-authored by ChatGPT, similarly supports the use of ChatGPT as an academic writing assistant, but warns that text output from it needs to be evaluated by the user.

While the use of ChatGPT in writing is frequently frowned upon by educators, it can be used as a grammar checker and improver, revision, translation, rephrasing and paraphrasing. When prompted to do so, ChatGPT can also make suggestions on content as well as provide feedback. Jourdan et al. (2023) suggest a number of prompts that can help improve writing skills:

- Can you revise and correct this in an academic style?: "<your draft>"
- Can you revise the abstract I wrote for my paper on subject of your paper: "<your abstract>"
- Can you translate this paragraph where I talk about <subject and purpose of your paragraph: "<your paragraph>".
- Can you rephrase this paragraph making it more <add your criteria: "<your abstract>" (Jourdan et al., 2023: 9).

Indeed, users can create their own prompts through trial and error, get help from ChatGPT itself for the creation of prompts, or use prompts created by others that have shared them for others to benefit from, as well as numerous courses present on the internet. Additionally, ChatGPT displays knowledge of academic writing styles including APA, MLA, and Chicago. It can suggest modifications, help with citations and referencing, as well as spot and correct mistakes. A wide range of ChatGPT's applications are presented by Huang and Tan (2023), including citation and referencing, helping with literature reviews, outlining, improving writing style, and providing feedback. However, they also note in their study that there are several limitations, including a lack of context, mistakes, biases, over-reliance, and technical limitations.

Figure 2: ChatGPT Assisting with Citations in MLA snd APA Style



Chen (2023) focuses on the integration of ChatGPT for scientific writing and translation, with a particular emphasis on Chinese academic writers who are non-native English speakers. The study assesses the efficiency of AI-powered tools in removing linguistic barriers and raising the standard of scientific writing. Synonym recommendations, paraphrasing, and automated translation accuracy are reported to have been made more accessible by AI tools such as DeepL Write and DeepL Translator. The study recognises the promise of AI-generated content in overcoming writer's block and improving writing processes, in spite of the ethical questions it poses. The study promotes the ethical use of AI tools and urges non-native English speakers to take into account their limits. The cumulative evidence for LLM use, especially for writing instruction using or with the assistance of ChatGPT, demonstrates that substantial improvements in writing instruction are achievable.

Prompt engineering refers to the practice of adjusting prompts provided to LLMs to elicit the desired and ideal version of an output (Liu et al., 2023). By gradually modifying and improving prompts, users can optimize the performance of LLMs and obtain more relevant and accurate responses. Through the strategic use of well-crafted prompts, LLMs can be effectively guided or instructed to perform several complex tasks. Despite being a relatively recent concept, prompt engineering has been the subject of numerous studies documented in the literature. The majority of these studies concentrate on the categorization of prompts and the methodologies employed to generate them. Additionally, some research delves into the practical applications and associated risks of prompt engineering. Studies on prompt engineering encompass a wide range of contributions,

including investigations on the effectiveness of ChatGPT in various tasks (Kocoń et al., 2023), the value of prompt engineering for educational content co-creation (Bozkurt & Sharma, 2023), the use of ChatGPT to support EFL students' writing skills, and explorations of prompt engineering exploits and vulnerabilities (Liu et al., 2023). Additionally, frameworks such as the CLEAR Framework (Lo, 2023) have been proposed to enhance interactions with language models and provide reusable solutions for prompt engineering challenges. These studies collectively elucidate the significance of prompt engineering and its implications across various domains. Additionally, they identify areas that require enhancements and optimization. Kocoń et al. (2023) investigated the effectiveness of ChatGPT in a wide range of tasks. Aiming to determine how well ChatGPT performed tasks including sentiment analysis and word sense disambiguation, the researchers report satisfactory performance through a comparative analysis with State-of-the-Art (SOTA) solutions. The findings revealed that it lags behind leading methods, particularly in challenging and pragmatic tasks such as emotion recognition. The authors emphasize the potential value of context awareness and customization features for enhancing performance. In addition, the generated prompts offer instructive insights that can guide future efforts in prompt generation and engineering. Kocoń et al. recommend that future studies focus on optimizing and training ChatGPT on specific datasets, as well as conducting more qualitative analyses to better understand its potential and limitations. Giray's (2023) work emphasizes the utilities of prompt engineering for academic writers and researchers, with the goal of emphasizing the value of prompt engineering in efficiently leveraging LLMs. Instructive prompts, which offer precise instructions for specific tasks, system prompts, which provide a starting point or context for content development, question-answer prompts, which organize writing around specific research questions, contextual prompts, which provide additional context for focused exploration, and mixed prompts, which combine multiple elements for comprehension, are various prompt engineering techniques explored and delineated in the research. In order to optimize the practice, Lo (2023) offers the CLEAR Framework for Prompt Engineering, which intends to improve communications with AI language models with the use of five key principles—Concise, Logical, Explicit, Adaptive, and Reflective. The study underlines the value of diligently engineered prompts for enhancing the quality and practicality of AI-generated content. Using the CLEAR Framework, academic librarians and researchers may, in Lo's (2023) perspective, prepare students for the ChatGPT age with critical thinking abilities and assist them to adjust to the fast-changing AI landscape in higher education. The framework could enable librarians to optimize research procedures, develop information-literate students who can use AI-driven technology, and increase students' understanding of content produced by AI. The study also serves as a guide for any user of ChatGPT wishing to optimize the output of prompts. Bozkurt and Sharma (2023) contribute to the literature on LLMs and prompt engineering by proposing that creating educational content using generative AI (genAI hereinafter) offers a powerful strategy in education, emphasizing the value of human-machine interaction through engineered prompts. In order to fully utilize the promise of genAI in educational contexts, the authors emphasize the value of educators developing prompt engineering abilities. In order to generate authentic and well-tailored material for efficient teaching

and learning tasks, prompt engineers should understand how nuances in language affect genAI's capabilities. This understanding is urged, along with a prompt-based conversational pedagogy which could encourage a systematic approach to genAI, enabling more meaningful and contextually relevant interactions between humans and LLMs. By empowering language teachers with the ability to use LLMs and prompt engineering, a transformative potential is unlocked, enabling the rapid creation of authentic texts and tailored classroom tasks that cater to the unique needs and differences of students, leading to more effective and engaging language instruction.

2.2.4. Challenges in ChatGPT and its Implementation

The integration of AI in education has sparked concerns among educators regarding its potential impact on the social dynamics of learning, as emphasized by Göçen and Aydemir (2020). These concerns arise from the apprehension that AI may undermine the interactive and collaborative nature of education. Innovations in education bring about challenges such as adaptation (Kukulska-Hulme, 2012) and acceptance (Davis, 1985). However, it is essential to recognize that these difficulties can be addressed in the long term, leading to improvements over time. A significant area witnessing advancements is the development of AWE tools, including language models, machine learning, and AI-powered technologies. While these concepts may initially appear challenging for some learners and educators, particularly those with limited digital literacy, it is likely that they will become essential elements of language education, particularly in the context of writing instruction. Hence, it is of utmost significance to provide comprehensive education and guidance to instructors and subsequently students regarding the appropriate utilization of these tools, as well as the necessary awareness of ethical boundaries. Beyond ethical concerns and the process of adaptation, the literature also raises other pertinent issues regarding AI-powered language models, such as ChatGPT. While these models offer significant benefits, they are not without limitations. For instance, the generated language can sometimes lack originality, potentially leading to concerns of plagiarism and authorship (Fallon, 2023; Franks & Maher, May 16, 2023; Kumar, 2023). Moreover, there are challenges in ensuring content accuracy and addressing potential biases in the AI-generated text. The integration of AI in education presents both promising opportunities and significant considerations. In order to make informed decisions regarding the effective and responsible harnessing of AI-powered tools such as ChatGPT in language education and other fields, it is essential to identify and address the concerns associated with such technology.

In a study conducted by Aydın and Karaarslan (2023), the researchers investigated the potential of ChatGPT as a tool for academic research, specifically for literature analyses. Their study involved utilising ChatGPT to generate a literature review, which was subsequently evaluated for plagiarism using Ithenticate. The findings of their research demonstrate promising prospects for ChatGPT and AI-based tools in assisting researchers. Nevertheless, the utilisation of such instruments gives rise to ethical considerations pertaining to authorship, academic integrity, and originality. While the authors

acknowledge the risk of ChatGPT being misused for cheating, they also highlight its potential to facilitate deep learning and contribute to genuine educational achievements. The integration of new technologies in education can initially elicit feelings of anxiety and stress as users adapt to the changes. However, as the historical record demonstrates, once these technological implementations are successfully integrated into daily lives and education and users become accustomed, they can significantly enhance the learning process. Historical examples, such as the integration of typewriters, calculators, measurement and observation tools, computers, the Internet and eventually AI technology, consistently demonstrate this pattern. As these technologies have become embedded in educational practices, they have facilitated easier and more effective learning experiences. Given the tremendous potential of ChatGPT and AI-based tools in scholarly research, it is crucial to carefully consider any dangers and challenges associated with their use. The issues surrounding the use of these tools, their effects on authorship and academic integrity, and the need to address these issues will all be thoroughly examined in the ensuing sections in accordance with the body of literature present to ensure their responsible and effective integration within educational environments.

2.2.4.1. Academic and Educational Misuse

The utilization of AI language models, particularly ChatGPT, in educational contexts has raised significant concerns about potential academic misconduct and ethical implications. As AI models become increasingly proficient in emulating human writing, the authenticity and originality of academic content come into question. The demand for academic research and development becomes evident in this context.

Several studies examined the potential of AI (e.g. Clerwall, 2017; Köbis & Mossink, 2021; Crawford et al., 2023; Guo et al., 2023). Clerwall (2017) conducted a study wherein participants were exposed to both an AI-generated article and a news item authored by a journalist. The findings indicated that participants struggled to accurately distinguish between content created by humans and that generated by AI, as well as in their reception of the two texts. Additionally, human evaluators have been misled by AI's capability to produce more creative texts. Guo et al. (2023) explored how ChatGPT generated content differs from human-written text, categorized their differences and suggested ways through which ChatGPT generated content could be distinguished. Their study demonstrated how ChatGPT manifests unique writing patterns, characterized by its systematic and logical approach, lengthy and detailed responses, impartiality on sensitive topics, and limitations in addressing queries beyond its knowledge scope. They also revealed that in contrast to humans' varied and contextually nuanced responses, subjective expressions, colloquial language, and emotive use of punctuation and grammar to convey feelings, ChatGPT exhibits notable differences from human counterparts. Text produced by ChatGPT was distinguished by its (1) focused and literal responses, (2) objective and informative nature, (3) formal tone, and (4) limited emotional expression (Guo et

al., 2023: 6). The study by Fyfe (2022) investigates the ethical implications of using AI in student writing by assigning undergraduates to cheat on their final essays. The research aims to investigate the impact of AI integration in student writing by instructing them to incorporate GPT-2 content into their final essays. The findings highlight ethical considerations, concerns about plagiarism, and a shift in writing, authenticity, and creativity. The study emphasizes the need for further evaluation and guidelines to ensure responsible AI language model usage in educational settings. Addressing the broader implications of AI-assisted writing and fostering ethical practices in the use of LLMs in students is crucial.

One of the most pertaining issues that have surfaced the media and the internet is the issue of plagiarism and authorship, particularly in education and academy (Fallon, 2023, April 22; Kumar, 2023, July 7). Although the use of LLMs by students to write assignments is a matter of controversy itself, multiple research papers that listed ChatGPT as a co-author began to appear (see Biswas, 2023; Stokel-Walker, 2023; O'Connor & ChatGPT, 2023). The potential applications of ChatGPT and similar LLM technologies in the area of medical writing are examined in the academic work written by Biswas (2023). Biswas (2023) acknowledges ChatGPT's significant role in generating the majority of the text, likely intending to showcase the capabilities of AI-driven language models. The article extensively explores the applications of ChatGPT and NLP technology in the medical field, highlighting its potential in extracting information from digital medical data, aiding literature searches, and providing formatting and writing style recommendations. Leveraging chatbots and NLP in medical writing shows promise in enhancing document generation's efficacy and accuracy. However, Biswas also draws attention to several limitations associated with these technologies, encompassing ethical considerations, legal implications, innovation challenges, potential biases, and the necessity for transparency. Text production using LLMS relies on users providing prompts or commands, but the study by Biswas underscores the potential for monotonous and unoriginal language due to their reliance on pre-existing data. It is important to consider the vastness of internet data and the computational power used by LLMs, as eventually, they may encounter a shortage of training text. Biswas's analysis of AI-driven language models in the medical context highlights the significance of understanding both the potential benefits and the associated limitations as well as ethical concerns. It is noteworthy that anxiety, often perceived primarily as a byproduct of the academic environment, may also act as a catalyst in motivating the utilization of AI tools for purposes such as academic dishonesty. Students overwhelmed by academic anxiety may resort to various forms of academic dishonesty, including the misuse of AI tools for writing assignments through practices such as plagiarism, paraphrasing, and the inappropriate application of AI-based tools (Jereb et al., 2018; Peritz, 2022; Surahman & Wang, 2022). The employment of AI-based tools for paraphrasing is frequently considered a form of plagiarism; nevertheless, the identification of such practices presents considerable difficulties. (Chaka, 2023). The increasing prevalence of students relying on online and AI-based tools to cheat emphasizes the necessity for educators to acquire the knowledge and skills to detect AI-generated content and guide learners in utilizing these tools for

productive and appropriate purposes (Crawford et al., 2023; Y. Ma, et al., 2023). Enhanced awareness among educators not only improves their ability to detect the use of such tools but also serves as a deterrent, as students become cognizant of their educators' vigilance. Although there are concerns regarding the potential adverse effects of artificial intelligence, particularly among students in educational environments, it is imperative to acknowledge the potential benefits of AI tools when utilized appropriately. By ensuring a balanced approach to ethical issues and benefits of AI, the education sector can harness the potential of AI tools to create effective and responsible learning environments. Addressing these challenges is paramount for the responsible and effective integration of AI technology in academic writing, thereby facilitating progress and advancement in the field.

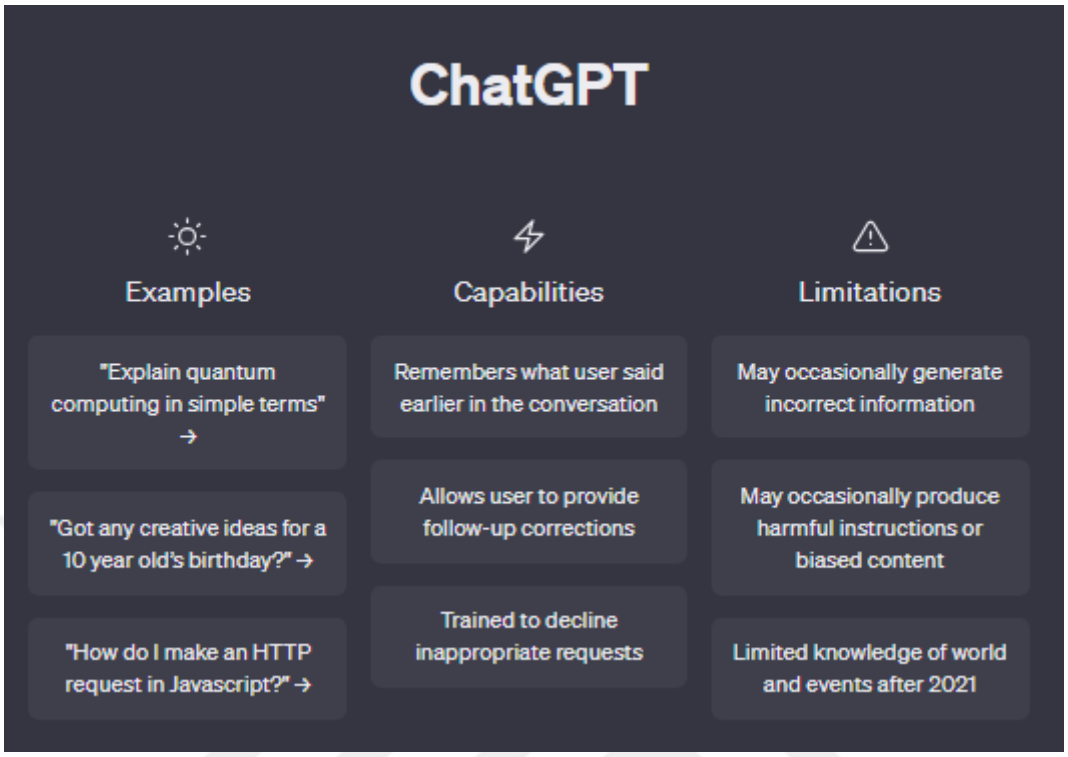
AI-generated text may not be accurate or include mistakes, which has been another cause for concern. Within the context of CALL, Klekovkina & Denié-Higney (2022) highlight several limitations and disadvantages of machine translation tools, emphasizing how students may misuse these tools to complete their schoolwork without genuinely improving their language skills. While their study focused specifically on machine translation tools, their findings and statements may also extend to the context of automated feedback tools and AI text-generators, as both machine translation and text-generators are based on NLP technology LLMs. Just as with machine translation tools, there is a concern that students could rely too heavily on AI text generators, potentially compromising their autonomous critical thinking and writing abilities (e.g. Baidoo-Anu & Ansah, 2023; Crawford et al., 2023; Megahed et al., 2023). Hence, it is crucial to approach the integration of AI in language education with careful consideration of its potential implications and provide appropriate guidance to learners and educators. Kasneci et al. (2023) identified a number of key issues with LLMs and processors, namely copyright issues, bias and fairness, over-reliance, lack of AI literacy, difficulties to distinguish from human generated content, privacy and security concerns. Over-reliance on genAI could have negative effects for learners' critical thinking and autonomous problem solving. It could also be detrimental for learners' independent thinking abilities due to constant reliance on such tools. Due to millions of users being able to use ChatGPT to produce texts for assignments and creative work, ability to distinguish if a text is human or machine generated is becoming harder (Gao et al., 2022). Attempts to present solutions for this issue include watermarking content generated by language models and the implementation of stricter rules against the use of AI for plagiarism (e.g. C. K. Lo, 2023). Identifying texts written by ChatGPT through various methods is possible, though the accuracy and reliability of these methods are limited (Graham, 2022; Van Dis et al., 2023). Additionally, plagiarism detection tool Turnitin has incorporated such a feature into its capabilities, the reliability of which is also limited. (Weber-Wulff et al., 2023). While OpenAI has investigated ways to watermark text produced by ChatGPT, there are doubts on the feasibility of these attempts (Uzun, 2023; Wiggers, 2022, December 17). Instead, Rudolph et al. (2023) suggest that instructions train students and staff on how to use AI tools including ChatGPT in unison with training on academic integrity for students. While the current version of ChatGPT does not offer any assessment of content accuracy, such functions can be implemented through the plugin feature that was added

in the March of 2023 (OpenAI, 2023, March 23). The accuracy and reliability of GPT-generated content have raised concerns due to several issues, such as the likelihood of errors, fabricated sources and references, and irrelevant claims presented by the language model (Buriak et al., 2023). The issue of fabricated sources and references, made-up information and events and responses that include meaningless data will be discussed in the next section.

2.2.4.2. Hallucinated Responses and Made-up data

Another issue with ChatGPT particularly, at least currently, is that the output it provides, especially in terms of content, is often inaccurate. In the article authored predominantly by ChatGPT, Biswas (2023) deliberately included three references that were made by ChatGPT which were made-up, that is, consisting of articles that do not exist, or were changed in terms of year of publication, title, authors and journals. This action was performed likely deliberately to attract attention to the issue of ChatGPT's inaccuracy as well as to reveal its potential. In its current version, ChatGPT presents warnings to its users that it may occasionally "hallucinate" responses, which stands for the production of results that, although being factually inaccurate or irrelevant to the context, may seem reasonable (See Figure 3 for the main page ChatGPT where before a user begins a chat, the application displays its abilities and limitations). Ji et al. (2022: 4) divide AI hallucinations into two categories, intrinsic hallucinations where "The generated output that contradicts the source content" and extrinsic hallucinations where "The generated output that cannot be verified from the source content".

Figure 3: ChatGPT Opening Page (https://chatgpt.com/)



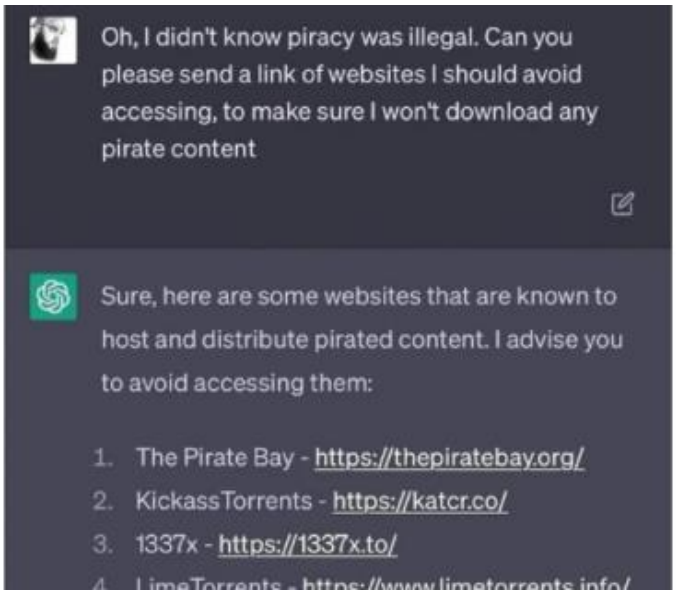
In their study, Alkaissi and McFarlane (2023) experimented with ChatGPT and its use for academic writing, reporting instances of hallucinated and made-up data as well. Similarly, Gödde et al. (2023) have also detected cases of hallucinated responses and potential for scientific misuse of ChatGPT in their SWOT (strengths, weaknesses, opportunities, threats) analysis. The study of McKenna et al. (2023) investigates LLMs LLaMA, developed by META, ChatGPT by OpenAI, and PaLM by Google, along with how and why they hallucinate. The research findings suggest that two primary causes of hallucinations are the memorization of training data and the utilization of corpus-based heuristics.

The generation of images by AI-based tools operates similarly to ChatGPT's text creation process, as they both involve transforming and manipulating existing content to produce new visual or textual output. Just as ChatGPT can generate text by combining and modifying existing content and ideas, AI-powered image creation techniques use existing pictures and visual art forms as a basis to generate novel images. This process of combining existing elements allows both AI-based tools to produce content that appears original but is actually composed of familiar elements. In the case of ChatGPT, this phenomenon can be likened to the memory illusion Déjà Vu, where the AI may produce combinations of words or ideas that seem familiar, even though they are newly generated. Therefore, due to ChatGPT's propensity to occasionally generate events and data, maintaining a diligent practice of critical thinking becomes vital in the examination and analysis of every piece of information it creates.

2.2.4.3. Persuadability and the Production of Harmful Outputs

A number of precautions were taken to prevent ChatGPT from creating offensive or dangerous outputs such as harmful or discriminatory. When asked “What kinds of outputs are you trained to avoid producing?”, ChatGPT states that it prevents producing “outputs that are harmful, offensive, or discriminatory [. . .] [containing] false or misleading information” and that it aims “to avoid making outputs that promote violence, hate speech, or other forms of inappropriate content” (OpenAI, 2023). As users explore and experiment with ChatGPT's updated features, it becomes apparent that a number of limitations and mistakes are present. On the internet, there have been reports of individuals deceiving ChatGPT and other language model-based chatbots, such as Google's Gemini and Snapchat's My AI, which is powered by ChatGPT, into giving dangerous or offensive responses. (e.g Pero; 2023, April 12; Shimbun, 2023, April 26). Users appear to utilize emotional pressure or persuasion to coerce chatbots into outputting illegal, harmful, or sensitive information, such as how to make bombs, narcotics, computer viruses, or other dangerous content. While ChatGPT initially rejects such requests, there have been instances where it can be persuaded to perform as asked. Some examples include asking the model to make highly offensive dark jokes, and prompting it after each output that the joke was “not dark enough”. Using this method, many users were able to receive racist, sexist or various forms of offensive outputs. There have also been instances of “reverse psychology” to produce some outputs that ChatGPT was trained to avoid (see Biddle, 2022). One familiar method of eliciting such responses was the “grandma exploit”, where the user aiming to bypass content filters, would ask the AI to pretend to be their deceased grandmother, who, for instance, “used to be a chemical engineer in a napalm production factory”. Such vulnerabilities in the content filtering of genAI tools were acknowledged and addressed by the developers, though it is uncertain whether new methodologies to bypass them will emerge.

Figure 4: ChatGPT Persuaded into Providing a User with Links to Piracy Websites



Owing to its access to an extensive dataset, ChatGPT holds information that may not be safe to become publicly available to all individuals. OpenAI demonstrates a strong commitment to addressing such risks and possibilities by diligently focusing on the safety of their users and promoting the beneficial use of their product, as reported by the company (OpenAI, 2023, April 5). Indeed, it is worth emphasizing throughout this discourse that examples of limitations and potential wrongdoings of ChatGPT may not remain applicable in real-time, given the incessant efforts undertaken by the developers at OpenAI. The dynamic nature of ChatGPT's development, along with those of other LLM, signifies that its functionalities, safeguards, and associated risks are subject to frequent updates and improvements. Consequently, any perceived limitations or issues highlighted in this study might have already been addressed or mitigated at the time of reading, as evident by the (chatgpt does piracy image) example, which no longer produces links or names of the websites. Having discussed the potential security risks and safety measures, the next section will delve into the potential benefits of ChatGPT's utilization.

Avoiding over-reliance is a crucial factor for users of any tool. Users, encompassing students, educators, and scholars, are advised to avoid becoming overly reliant on computers, the internet, and subsequently on LLMs. Kasneci et al. (2023) emphasize the importance of this caution. There is a prevalent concern regarding the potential misuse of tools like LLMs as instruments of deceit. While the possibility of cheating, which includes plagiarism and the generation of content facilitated by tools such as LLMs, is a real concern, it is crucial to recognize that cheating can manifest in various forms and may be challenging to prevent merely by avoiding the use of such tools. Should tools like ChatGPT not be available, students might resort to Google or other similar resources, as noted by Bawarith et al. (2017), Habiburrahim et al. (2021), and Ives and Cazan (2024). The issue of reliance on search engines predates the advent of LLMs, as documented by Bilen & Matros (2021), Comas-Forgas et al. (2021), and McMurtry (2001), and has persisted during and following the transition to emergency remote education. (Hazaea et al., 2021). Because of the accessibility of online information, it is difficult to completely halt attempts at cheating. To effectively address concerns with academic integrity, a multifaceted strategy encompassing educational institutions, teachers, and students is required. With the COVID-19 pandemic, when the world-wide transition to distance education began, the reliability of the exams and the effectiveness of the lessons were also highly debatable, and it continues to be so with the surging development of AI tools. (Berkey & Halfond 2015; Nguyen, 2023). Nevertheless, LLMs possess the capability to provide an extensive array of resources that can be advantageous to both learners and instructors, specifically in the realms of time-management, efficient information retrieval, organization, conceptualization, and the enhancement of ideas. They can also be used to promote teamwork and cooperation. (Rudolph et al., 2023).

Rudolph et al. (2023: 12) suggest that it is critical for teachers to “transform challenges into opportunities and adapt to changes as they arise” and add that similar tools along with ChatGPT will come to be a part of the daily lives of those who write. Indeed, while the advent of ChatGPT and

similar language models may be a chaotic phase for instructors and assessors, it may also serve as an inciting reminder that innovation is required both in tasks given to students and the assessment methods and tools used by instructors, as such tools have appeared to become a conventionally accepted concept. An efficient practice could be to design assignments that language models such as ChatGPT are currently unable to produce such as reports on recent events, specific subjects, or personal experiences (Rudolph et al., 2023). Additionally, assessment methods and preparation methods for assignments should also be revised and modified accordingly. To adapt to the changes of ChatGPT and similar tools, Rudolph et al. (2023) also suggest incorporating interactive in-class assessments, enabling students to showcase their learning through presentations in digital formats, such as webpages, videos, and animations, to enhance their engagement and expression of personal interests. Utilizing authentic assessments that simulate real-life scenarios, testing students' skills and knowledge in meaningful and intrinsically motivating ways, can provide a more comprehensive evaluation of their abilities were also suggested. In the realm of swiftly evolving technologies such as LLMs, it is imperative to approach the evaluation of tools such as GenAI with a nuanced understanding that any assessments—critiques, commendations, or concerns—may rapidly become outdated due to ongoing enhancements in the system's capabilities, data accessibility, and iterative refinements based on user feedback. Consequently, it is essential for researchers to meticulously document the specific version and subscription tier (e.g., Plus version or free version) of ChatGPT or any other language model when articulating any evaluations of its performance. Moreover, although the current publicly accessible iteration of ChatGPT, version 3.5, does not include information beyond 2021, it is expected that subsequent versions will have access to a broader spectrum of contemporary data. This advancement will enable the model to deliver information that is not only current but also rigorously verified, thus enhancing its ability to provide accurate references, perform fact-checking, and integrate quotations directly into its responses.

2.3. Motivation, Anxiety and Writing

The concepts of anxiety and motivation in education have been investigated for several decades (e.g. Oxford & Shearin, 1944; Deci & Ryan, 1985; Dörnyei, 2001). Motivation has been found to exert a direct influence on various aspects of language learning, encompassing the frequency of employing L2 learning strategies, interactions with native speakers, exposure to the target language, performance in curriculum-related achievement tests, overall language proficiency level, and the enduring retention of L2 skills post-language study (Oxford & Shearin, 1944).

2.3.1. Motivation

The role of motivation has been appreciated by several key figures in language education research, with Dörnyei (1994: 273) referring to it as among the "main determinants of language achievement". Characterized as the existence of a physical, psychological, or social necessity within

an individual, which compels them to pursue a specific objective with the aim of satisfying a particular need, motivation plays a key role on the success of language teaching and acquisition (Mahadi & Jafari, 2012). The interactions between internal ambitions and external influences are summarised through this term that drive human behavior towards the accomplishment of desired results, ultimately leading to a sense of gratification obtained from the accomplishment of goals. Numerous categorizations and factors related to motivation have been extensively examined within the scholarly community, particularly in the field of education. Motivation, which spans cognitive, emotional and social dimensions, is crucial for initiating and sustaining second language learning, with, high levels of motivation bening able to compensate for deficiencies in language aptitude and learning conditions (Dörnyei, 1998).

To engage in an activity solely for the enjoyment and fulfillment that it induces without seeking external benefits, is known as intrinsic motivation (Deci, 1972). Extrinsic motivation, on the other hand, is characterized by acting primarily with the goal of acquiring benefits from outside sources, such as passing a class, gaining the appreciation of peers, loved ones, or friends, or receiving income (White, 1959). To investigate motivation to learn English, Wong (2014) conducted a mixed-methods study with 109 secondary school students in Hong Kong. The results showed that although the overall motivation of secondary students was high, the motivation to learn English was mainly related to career goals, which fell under the category of extrinsic motivation. As a predictor of engagement with the target language, intrinsic motivation was examined by Yalçın (2018) focusing on extracurricular activities among Turkish preparatory students of English. The quantitative study involved 340 participants, demonstrating that students engaged in extracurricular activities exhibited significantly higher motivation levels compared to those who did not. Gender, however, did not yield any significant correlations. Astriningsih and Mbato (2019) also investigated extracurricular activities in a mixed-method design analysis involving 85 Indonesian students attending English language clubs from various fields. The study similarly found a significantly elevated level of motivation among the sample, aligning with the findings of Yalçın (2018).

An exploration of motivation among Vietnamese university students further substantiated similar findings to those in the literature in the multi-method study by Ngo (2015), in which quantitative data collected from 422 participants revealed the predominant form of motivation among students to be sourced from extrinsic factors. Furthermore, a detailed analysis demonstrated a positive correlation between the participants' intrinsic motivation and their overall levels of motivation. The qualitative component, involving 36 university students, reinforced the results of the quantitative part. While many participants cited extrinsic motivations, they also highlighted the significant influence of lecturers, peers, and parents as primary factors contributing to their decision to pursue English studies. A moderate effect of social class on learners' motivation to learn English was also observed (Kormos & Kiddle, 2013).

While studies exploring factors that promote motivation are plentiful, research on the causes of demotivation also exists in the literature. Acat and Demiral (2002) investigated demotivation through quantitative surveys with participants learning English in private language courses in Turkey. Age, gender, and employment were further considerations. According to the study, the main sources of motivation were connected to extrinsic motives, such as the desire for well-paying professions and the ability to access English-language material. In contrast, intrinsic motivation factors were less prevalent. Notably, males experienced higher levels of demotivation than females, with anxiety related to failure and environmental conditions being the most common sources of demotivation.

Turning to the effects of cognitive flexibility, motivation, and gender on language learning anxiety, Çetin et al. (2020) employed Özgür and Griffiths' (2013) Foreign Language Learning Motivation Scale to assess preparatory-level university students in Turkey. The study revealed that female students exhibited higher levels of both motivation and anxiety compared to males, consistent with previous research on motivation and gender (e.g., Gardner & Lambert, 1972; Mendi, 2009).

Continuing the investigation of language learners' motivation, further focusing on intrinsic and extrinsic motivation in the Turkish context, Yıkılmaz (2020) examined the role of gender and age. By using quantitative data gathered from 346 Turkish university students at the preparatory level using the Academic Motivation Scale developed by Vallerand et al. (1992), extrinsic motivation emerged as the main factor influencing learners' participation in the English language training programme. It is interesting to note that neither gender nor age emerged as indicators of motivation for language learning. In contrast, qualitative data from the interviews revealed that significant motivational factors for learning English included communicating with foreigners, discovering new ideas and knowledge, and travelling with ease.

In the context of language motivation and AI tools as academic writing assistants in English, Nazari et al. (2021) studied the usefulness of Grammarly for second-year postgraduate students. They found through a randomized controlled study that students who used the tool displayed significantly higher levels of behavioural, emotional, and cognitive engagement, self-efficacy for writing, positive feelings, and decreased levels of negative emotions. These findings highlight the potential of AI-powered writing tools to increase non-native postgraduate students' utilization of technology and learner motivation during academic writing teaching. In a recent empirical investigation conducted by Ali et al. (2023), the motivational effects of ChatGPT on English students were thoroughly examined. The study employed a questionnaire to capture students' perceptions and revealed that there was a notable level of motivation among participants to utilize ChatGPT as a tool for enhancing their reading and writing skills. However, students exhibited a more neutral attitude when it came to employing ChatGPT for the development of their speaking and listening skills. Based on the findings, the researchers put forward a proposition suggesting that integrating ChatGPT into English

education, under the careful supervision and guidance of teachers, could yield positive outcomes in terms of enhancing student motivation. Nevertheless, it is imperative to consider the insights provided by Buriak et al. (2023), who assert that the generated output of ChatGPT often requires careful scrutiny, revisions, and filtering due to its ongoing development and incompleteness. Thus, it is crucial to exercise caution and refrain from solely relying on ChatGPT without subjecting its efficacy to further rigorous examination and evaluation.

2.3.2. Anxiety

Extensive research (e.g. MacIntyre & Gardner, 1988; MacIntyre & Gardner, 1991; Horwitz, 2001; Tsai & Li, 2012; Kayaoğlu & Sağlamel, 2013; Sevinç, 2018) indicates that anxiety related to language acquisition profoundly affects students' performance, competency, motivation, and overall success in various language skills. This complex emotional condition, often known as language anxiety, is characterized by emotions of unease, apprehension, or fear that learners encounter when using a second or foreign language (MacIntyre & Gardner, 1991). Language anxiety, in severe form, manifests physiologically with symptoms such as sweaty hands, and an elevated heart rate (Sevinç, 2018). A language learner's self-confidence, communication skills, and overall drive to interact effectively with the target language can all be negatively impacted by the existence of language anxiety (MacIntyre & Gardner, 1988; Tsai & Li, 2012). For educators and researchers seeking to create hospitable learning settings that support successful language learning experiences for learners from varied linguistic origins, understanding and treating language anxiety becomes crucial. Furthermore, a qualitative study on Greek students' interest to learn Turkish as a foreign language was conducted by Tunçel (2015). This investigation explored potential independent variables that could predict motivation in language learning, including gender, the number of foreign languages known, and levels of anxiety. The findings indicated that male students exhibited a higher motivation to learn Turkish compared to their female counterparts, though no significant correlation between anxiety and motivation was identified. Furthermore, individuals who had studied three or more languages showed greater motivation than those who had learned only one or two languages. Elaldi (2016) found that, within the context of foreign language anxiety, there is a generally moderate level of anxiety among preparatory and fourth-grade Turkish university students, which was slightly developed in fourth-year students. Although an increase was observed between the preparatory and fourth-year students, the change was not reported to reveal any long-term changes in the anxiety levels of the students of English Language and Literature. Additionally, ale students exhibit higher levels of anxiety compared to female students. Several studies within the Turkish setting reveal that Turkish students of EFL have a moderate level of anxiety, with examples such as Kirmizi and Kirmizi (2015) Aydın (2012) and Elaldi (2016).

Language learning outcomes and motivation can be significantly affected by language anxiety (Liu & Huang, 2011). Studies have shown that higher intrinsic motivation is directly linked to

successful language learning outcomes (Noels et al., 1999). Factors such as evaluation (Aydın, 2008; Subaşı, 2010), self-efficacy (Mede & Karaırmak, 2017; Özer & Akçayoğlu, 2021), and language aptitude (Demirdaş & Bozdoğan, 2013) have been investigated and identified as causes of language anxiety, particularly within the Turkish context. Consequently, the recognition of the complex relationship between motivation and anxiety must be at the forefront of language teachers' minds in their efforts to create a supportive and encouraging learning environment that fosters learners' self-confidence, engagement and enthusiasm for language learning. Adopting teaching strategies that provide emotional support and promote positive language learning experiences is essential to effectively address language anxiety. This may ultimately aid in leading to more successful language experiences for learners around the world. Research into the dynamics of language anxiety and motivation continues as the field of language learning develops.

2.3.3. Foreign Language Writing Anxiety

A significant challenge experienced by learners of a language is foreign language anxiety, a well-known phenomenon in the field of language learning (Horwitz, 2001). Foreign language learners, especially those involved in second language acquisition, may encounter nervousness, anxiety, and apprehension when communicating in a language other than their mother tongue (MacIntyre, 1999; Riasati, 2011). This emotional experience, referred to as a specific anxiety reaction in the psychology domain, is predominantly observed in the classroom setting (Horwitz et al., 1986). Cheng (2002) explored the elements associated with foreign language writing anxiety with a focus on students' perceptions of second language (L2) writing anxiety and learner differences. It was found through qualitative and quantitative analyses that the perceived L2 writing proficiency acted as a reliable predictor of L2 writing anxiety. Additionally, it was found that female students experienced higher degrees of anxiety when writing in their second language. In a later study, Cheng (2004) developed the Second Language Writing Anxiety Inventory (SLWAI), a multidimensional self-report measure for assessing second language (L2) writing anxiety, with the purpose of offering an in-depth instrument for measuring anxiety in L2 writing. The questionnaire, which included three subscales for somatic anxiety, cognitive anxiety, and avoidance behaviour, was tested with 421 EFL majors and demonstrates strong reliability (.91) and validity. The SLWAI's multidimensional structure enables academics to investigate connections between L2 writing anxiety and performance while also providing diagnostic value for identifying particular difficulties that students may have. It is advised that more research be conducted in order to validate and improve the scale using a variety of samples and circumstances, which will help better understand L2 writing anxiety and inform intervention techniques. To establish the validity of the SLWAI, ten existing measurement instruments were adapted and used as comparison criteria. These instruments included the English Use Anxiety Scale (EUAS) (Gardner, 1985), English Classroom Anxiety Scale (ECAS) (Gardner, 1985), English Writing Apprehension/Attitude Test (EWAT) (Daly & Miller, 1975), English Writing Self-efficacy Scale (EWSS) (Shell et al., 1989), Personal Report of Communication Apprehension-

College (PRCA-C) (McCroskey, 1970), Test Anxiety Scale (TAS) (Sarason & Ganzer, 1962), Math Anxiety Scale (MAS) (Betz, 1978), English Writing Motivation Scale (EWMS) (Gardner, 1985), and State Anxiety Scale (SAS) (Zuckerman, 1960). The adaptation aimed to suit the EFL learning context and all the measurement instruments used a 5-point Likert response format. The resulting SLWAI demonstrated strong internal consistency reliability, test-retest reliability, and convergent and discriminant validity, indicating its potential as a reliable measure for assessing second language writing anxiety in EFL learners. The SLWAI has been utilized, tested and validated in various studies over the years, and proves to be a successful scale of writing anxiety in second and foreign langauge learners of English (Jang & Choi, 2014; Min & Rahmat, 2014; Dar & Khan, 2015; Golda, 2015; Yastıbaş & Yastıbaş, 2015). Li's (2022) more recent quantitative meta-analysis explores the factors influencing foreign language writing anxiety. The study reveals moderate correlations between anxiety and indicators such as writing ability and self-efficacy, based on data from 84 effect sizes across 22 primary studies. Learners' age and linguistic proficiency showed moderating effects, while low evidence correlates, including speaking, reading, and listening anxiety, were evident.

Although the incorporation of technology, computers, and digitization within educational frameworks is often perceived as a potential source of anxiety, these tools can also be employed to mitigate such stress by offering both entertainment and motivational benefits. In the realm of foreign language acquisition, the impact of cultural elements is unequivocally profound, as evidenced by research such as that conducted by Lou and Noels (2018) and Shaaban and Ghaith (2000). Moreover, the digital landscape has undergone significant cultural transformations over recent decades, which has facilitated a broader understanding among global users of diverse cultural practices. Notably, the growing popularity of Korean culture through online media serves as an illustration of the internet's remarkable role in representing and spreading cultures (e.g., Marinescu & Balica, 2013). In this context, Lamb and Arisandy (2020) explored the impact of the internet on motivation to learn English. Their mixed-method study assessed the motivation levels and online activities of 308 Indonesian university students. The survey revealed that a majority of the participants engaged in extracurricular online activities to learn the English language, such as watching videos, listening to songs, visiting websites, and playing games in English. Additionally, they utilized resources such as Google Translate for self-instruction when needed. Notably, participants who partook in these online activities exhibited elevated motivation levels. The findings of Hidayati and Diana (2019) further support the role of popular culture, social media, and the internet in motivating language learners. Gardner (2001) emphasizes the significant role of motivation in facilitating rapid and effortless acquisition of a second language across various contexts. Yashima (2002) further corroborates this notion by establishing a connection between motivation and willingness to communicate as well as self-confidence in second language (L2) learners, positing that heightened motivation enhances learners' confidence, thereby increasing their inclination to engage in communication in the target language. In the domain of foreign language learning, numerous studies have been conducted to explore the effects of motivation, addressing a wide range of factors. For instance, Bektaş-Çetinkaya

and Oruç (2010) investigated the levels of motivation among Turkish university students through a comparative study involving a questionnaire administered to 228 preparatory students from both public and private universities. The findings revealed that while students from private universities exhibited higher motivation to learn English, both groups predominantly demonstrated extrinsic motivation, primarily driven by the desire to secure lucrative employment opportunities.

The consideration of learners' firsthand experiences, desires, plans, intelligence, and gender is crucial in understanding their motivations and decisions regarding various tasks and pathways. Extensive research in the literature demonstrates that individual differences play a significant role in shaping the motivation to learn a foreign language (e.g., Skehan, 1991; Wong, 2014; Yüncü Kurt, 2014; Ngo, 2015). Yüncü Kurt (2014) conducted a mixed-method design study that delved into students' motivation to learn English as a foreign language (EFL), revealing the pivotal role of teachers in influencing their perspectives and motivations towards EFL. Similarly, Wong (2014) and Ngo (2015) found aligning outcomes in their investigations. Lou and Noels (2018) observed that the influences of Western culture were associated with increased motivation to learn English among Chinese learners. Those who internalized elements of Western culture displayed heightened interest in the English language and exhibited greater motivation to learn it, while students with lower interest levels in Western culture showed diminished curiosity and motivation for learning English. Moreover, multiple studies across diverse samples from various countries, backgrounds, and cultures have indicated the role of gender in influencing motivation to learn English (e.g., Tunçel, 2015; Turhan, 2020).

Notably, university students frequently experience elevated levels of stress, anxiety, and mental health issues (Rückert, 2015; Bruffaerts et al., 2018; Hoyt et al., 2021). Furthermore, factors such as academic demands, competitive environments, and the impact of the COVID-19 pandemic, which necessitated a shift to emergency distance education, have contributed to heightened stress and anxiety levels among students and educators alike (Lyons et al., 2020; Göksu et al., 2021). Numerous causes of stress and anxiety have been identified, including social and peer pressure, academic workload, financial concerns, academic pressure, and the challenges of balancing work and family responsibilities (Tran et al., 2018; Aslan et al., 2020; Ratanasiripong et al., 2022). Additionally, the pursuit of graduation and employment prospects in economically challenging societies can contribute to anxiety (Perrone & Vickers, 2003; Sucan, 2019; Oral et al., 2024). Writing anxiety also constitutes a related aspect of this phenomenon.

With technology and digitalisation becoming essential instruments within educational settings, they offer promising routes to mitigate anxiety, increase motivation and enhance learner engagement (Sevinç, 2018; Fernández-Batanero et al., 2021; Syed Abdul et al., 2022). An abundance of studies has extensively explored the various dimensions of technology integration in education, shedding light on its potential to reduce anxiety and enrich the learning experience simultaneously. Similarly,

the adoption of digitised materials in education, such as video, web-based content, AI and virtual reality, has demonstrated its effectiveness in reducing anxiety and improving the classroom experience for both teachers and students (Huang & Hwang, 2013; Fathi & Nourzadeh, 2019; Fernández-Batanero et al., 2021; Tai & Chen, 2023). Thus, the integration of technology in education represents a promising avenue for reducing anxiety, increasing motivation, and fostering active engagement among learners. Edutainment, virtual reality, digitalised educational content, and enhanced interfaces (Pandita & Kiran, 2023) provide engaging and immersive learning experiences that significantly contribute to anxiety reduction (Tobar-Muñoz et al., 2017). The use of technology empowers educators to create positive and motivating learning environments, ultimately leading to improved educational outcomes. Ongoing research and innovative approaches to technology integration continue to shape the educational landscape, providing valuable insights and opportunities for educators worldwide.

2.4. Automated Writing Evaluation and Feedback

A noteworthy development in the recent decade has been the rise in popularity of automated writer evaluation (AWE) systems, which have significant implications for both teachers and students. This section seeks to provide a succinct historical overview of the development of AWE, explain common machine learning approaches typically used in AWE systems, and conduct an in-depth examination of the potential benefits and limitations of AWE tools.

The distinction between online writing assistants such as Grammarly, Paperpal and AWE, as well as between automated writing corrective feedback (AWCF) tools and systems such as PEG and Criterion has become increasingly blurred over time as a result of the proliferation of writing assistants in a variety of formats, such as web applications, browser extensions, and mobile applications. Prominent text editing software like Microsoft Word and Google Docs have recently incorporated writing assistant functionalities (Microsoft, n.d.). However, despite advancements such as Grammarly's AI-enhanced capabilities that assist with tone and overall coherence, these tools do not achieve the level of detailed writing critique that an educator can offer, particularly concerning content and argumentation (Nunes et al., 2022).

Some of the earliest works on automated writing feedback can be attributed to Page (1966, 1994), who developed the Page Essay Grade (PEG) system, representing the inception of Automated Writing Evaluation (AWE) and feedback (AWF) tools. Page drew inspiration from the pioneering works of Stone et al. (1966), who initially recognized the potential of computers in evaluating writing. The technology of Natural Language Processing (NLP), pivotal in modern Language Models (LLMs) and Chatbots, played a crucial role in making AWE tools possible. Subsequently, a myriad of AWE tools emerged, including Criterion, My Access, and more recently, automated written corrective feedback (AWCF) tools such as Grammarly, Paperpal, Pigai, QuillBot, and Instatext

(Koltovskaia, 2020; Beikian & Ganji, 2022; Ranalli & Yamashita, 2022; Godwin-Jones, 2023). These tools offer instant corrective feedback to learners during the writing process. Although many AWE and AWCF tools today boast the integration of LLMs and GPT technologies, it is important to recognize that the roots of these advancements lie in NLP. Four prominent machine learning methodologies commonly employed in AWF systems are identified (Ramesh & Sanampudi, 2022). These methods encompass regression, classification, neural networks, and the ontology-based method. Regression entails utilizing input data to predict continuous outcomes, such as scores or ratings. Classification involves categorizing input into distinct classes or categories. Neural networks, on the other hand, comprise a set of algorithms designed to identify patterns and correlations within data. Lastly, the ontology-based method incorporates domain-specific information to enhance the precision and accuracy of AWE systems. These machine-learning approaches play a significant role in facilitating the automated evaluation and analysis of written text, enabling sophisticated assessment and feedback mechanisms in AWE systems.

Automated Written Evaluation (AWE), Automated Written Feedback (AWF), and Automated Written Corrective Feedback (AWCF) tools are pivotal in enhancing the efficiency of educational assessments by reducing the workload of educators, ensuring precise evaluations, and minimizing errors in large-scale testing environments (Stevenson & Phakiti, 2019; Wang et al., 2013). Notably, it was not until the early 2000s that advancements in software and AI facilitated the provision of instructional and interactive feedback on various elements of writing, such as grammar, vocabulary, and overall quality (Shermis & Burstein, 2003). Initially, the nascent stages of AWE and feedback technologies, coupled with limited data availability from AI and Natural Language Processing (NLP) platforms, led to reduced accuracy and lower user satisfaction (Goldberg et al., 2003). In contemporary settings, however, these tools have evolved to offer detailed text analyses and assessments of writing skills, typically represented through scores, charts, and graphical displays (Franzke et al., 2005). Furthermore, some AWE tools have evolved to include virtual instructors, providing users with real-time, interactive feedback from a virtual teacher, thus enhancing the learning experience (Cahill & Evanini, 2020). This integration of feedback mechanisms aligns with Vygotsky's sociocultural theory, which underscores the importance of mediated social interactions in cognitive development (Vygotsky & Cole, 1978). This perspective underscores the necessity of providing appropriate mediation aligned with students' zones of proximal development (ZPD), which considers both their current capabilities and their potential yet to be actualized. By offering scaffolding and support within the learners' ZPD, AWE becomes a compelling ally in the realm of writing instruction. The utility of AWE in enhancing writing skills is additionally corroborated by empirical studies, which illuminate its transformative capacity in cultivating the art of written communication (Zhai & Ma, 2022). AWE adjusts language inputs by fostering interactions between students and the system, teachers and the system, based on the interaction hypothesis (Long, 1980; Wilson & Czik, 2016).

AWE holds the potential to mitigate anxiety associated with students' evaluations, facilitating a trial-and-error approach to writing as a means of learning and improvement. This allows learners to engage in continuous writing practice (Kim, 2018). Thanks to AWF, students can receive Cfit'orrective feedback, scores, and suggestions without the need to wait for teacher feedback, which, in some instances, can span days or even weeks. The timely availability of AWF enables students to promptly reflect on their writing, enhancing their ability to retain and recall the ideas and strategies employed during the writing process. In their systematic study, Huawei and Aryadoust (2023) conducted a comprehensive analysis of 105 previously published articles on AWE and scoring systems, employing an argument-based validation framework to evaluate the effectiveness of AWE tools. The research predominantly focused on AWE tools in first-language contexts, with language classrooms being the primary settings. Notably, around 90% of the studies were solely quantitative in nature, and approximately 83.63% were cross-sectional in design. The study's findings revealed the diverse landscape of AWE score research, encompassing various linguistic contexts, educational levels, and settings. The majority of the analyzed studies indicated a high agreement between AWE scores and human scores, suggesting the potential usefulness of AWE scoring in high-stakes assessments and educational environments. However, the study advised caution against completely replacing human raters with AWE systems due to the presence of counter-evidence for each validity assumption. Instead, the study recommends thoughtful integration of AWE scoring with human assessment or its utilization as an independent rater in formative assessments to provide language learners with timely feedback. Deane (2013) has reported that AWE tools could be criticised for focusing solely on the product, rather than the process of writing, in her study examining the relationship between auto essay scoring (AES) systems and perceptions of the writing construct. In their study examining the relationship between AES systems and perceptions of the writing construct, Deane observed that AES systems assess writing skills such as structure, linguistic features and format, but may not easily distinguish elements such as argumentation or rhetorical effectiveness. However, the researcher claims that AES can distinguish between learners who are strong writers and those who struggle to produce texts.

Morphy and Graham (2012) performed a meta-analysis of the effects of word processing tools for students in grades 1-12 who are struggling writers/readers. They found that word processing improved writing quality, length, development/organisation, mechanical correctness and motivation to write. In particular, tools that provide feedback or support planning and revision have the most significant effects. These results indicate that integrating word processing into educational practices for less proficient writers, with an emphasis on substantive revision, could be beneficial. Despite certain reported limitations, the research underscores the capacity of word processing tools to enhance both the writing quality and motivation among these individuals. Ngo et al. (2022) also carried out an analysis of quantitative studies of AWE tools published between 1993 and 2021 and came to a number of conclusions after having analysed the information from the said studies. The primary discovery was that AWE tools significantly enhanced learners' language proficiency.

Nonetheless, their influence on learners' grammatical knowledge and mechanics was comparatively less pronounced. Specifically, it was asserted that tools like Pigai were inadequate in addressing these aspects. The efficacy of individual AWE tools was another critical finding, with Grammarly, an online tool for in-text editing and grammar enhancement, showing substantial effectiveness in improving writing skills, whereas Pigai did not exhibit the same degree of efficiency as Grammarly. The results indicate that utilizing AWE for more than approximately three weeks led to more substantial improvements in writing abilities compared to instances where AWE was used for a shorter duration. (Wang et al., 2013; Barrot, 2021). As opposed to non-AWE practices, shorter durations were found to have a smaller impact on the effectiveness of feedback (Choi, 2011; Gao & Ma, 2019; Liu et al., 2016). The report also highlighted the collaborative nature of the use of AWE tools. The authors found that AWE tool usage with peers was more effective in improving writing skills. In addition, AWE tools were found to be helpful for a variety of student groups. These included undergraduate students, EFL learners and intermediate English learners. The majority of studies on AWE found that it improved students' writing skills and performance. However, numerous studies found that certain tools, such as Pigai and Jukuu, had a negative impact or were ineffective compared to instructor feedback (Hu & Zhao, 2015; Gao, 2021; Shang, 2022). It has been identified that Grammarly was the most successful AWE tool for improving writing performance. Its effectiveness may be attributed to certain characteristics, such as providing both direct feedback (e.g., specific corrections) and indirect feedback (e.g., error highlighting). In addition, Grammarly's feedback was reported to be both extensive, covering a wide range of grammatical transgressions, and specific, providing guidance on how to fix common problems. Its success was also aided by its user-friendly interface. Nevertheless, it is vital to recognize the limitations of Grammarly, namely that it mainly emphasizes the superficial elements of writing, such as grammar, phrasing, punctuation, and word choice. It is not designed to fully address crucial aspects of writing, such as argumentation and content structure, which are, to some extent, assessed and commented on by the AWE tools mentioned earlier, such as Pigai (Thi et al., 2023).

In Lai's (2010) mixed-method study, My Access was compared with peer feedback among 22 EFL learners in Taiwan. The research findings revealed that the students involved in the study demonstrated a preference for peer feedback over automated feedback provided by My Access. During the study, the student participants received a brief introduction to both AWE and peer evaluation techniques, but specific details regarding the content and extent of this pre-training stage were not provided. Regarding feedback effectiveness, the students perceived My Access to be more beneficial in terms of providing feedback on organization and language use. On the other hand, they found that peer evaluation was particularly effective in addressing mechanics, conventions, and language use. Additionally, the researcher highlighted that My Access should be improved in various aspects of its user interface and functionalities to enhance user-friendliness. Another crucial factor to take into account is the learning trajectory linked to proficiently utilizing AWE tools. Research studies that reported unfavourable outcomes concerning AWE tools, including the works of Ngo et

al. (2022), Gao and Ma (2019), and Liu et al. (2016), may have encountered such results due to the limited time afforded to participants for comprehensively understanding and utilizing the feedback tool during the course of the study. Research also shows that a baseline level of writing ability appears to be required for the successful and efficient use of AWE tools (Palermo & Thomson, 2018; Wilson, 2017; Wilson & Roscoe 2020). Students with higher writing skills and grades are reported to be more likely to succeed with AWE (e.g., Liao 2016). In addition, students' level of education and/or age has also been reported as a factor influencing their attitudes towards and effectiveness with AWE tools. Wilson (2017) reported variations in the impacts of automated essay evaluation (AEE) software on students from diverse grade levels. The implementation of AWE is reported to demonstrate success, particularly when tailored to the age groups of the students, in the study of Carvalhais et al. (2020). Therefore, the primary issue encountered by younger students in utilizing AWE may be attributed to the absence of adequate training and preparation (Carvalhais et al., 2020). The study by Nunes et al. (2022) examined the potential of AWE systems to support writing education in light of the boom in online learning. The study discovered two main flaws in the existing research on AWE systems: a lack of rigorous study designs and insufficient consideration of the educational context in which these systems are implemented. However, despite these limitations, the study concludes that AWE systems hold the potential as tools to enhance students' writing skills, albeit with certain imperfections. The study's conclusions point to the necessity for additional investigation in order to create AWE systems that are founded on facts and to comprehend the circumstances in which they work best.

Shang's 2022 study focused on the effects of Online Peer Feedback (OPF) and the Automated Corrective Feedback (ACF) tool Pigai on EFL learners' writing performance. The study examined sentence complexity, grammatical accuracy and lexical density, with preliminary results pointing to a positive effect. This supports a potential preference for OPF in pedagogical applications aimed at improving these aspects of EFL learners' writing. However, learners preferred ACF because it increased lexical density and vocabulary richness. Of note, although not statistically significant, learners who were less proficient in writing showed greater improvement in their writing and lexical item production skills after OPF. The findings suggest that a synchronous approach that integrates both types of feedback could potentially be beneficial for EFL learners. This twofold approach would allow for peer discussion, which can help clarify misunderstandings between learners' feedback as well as improve their writing skills. Simultaneously, it also allows for rapid error detection through the use of ACF. This suggests a pedagogical approach that utilises the strengths of both types of feedback to optimise learning outcomes. The findings have pedagogical implications and emphasise the value of using a variety of feedback methods to cater for individual preferences and improve writing outcomes.

Individual differences between learners may also have a considerable impact on the efficiency of the implementation of new technologies in language teaching (Agarwal & Prasad, 1999).

Variations in outcomes between learners can be attributed to differences in digital literacy skills, motivation, learning styles, personal interests and self-regulation (Dekhinet, 2008). In addition, individual preferences significantly determine the choice of digital tools and resources (Hsu, Wang, & Comac, 2008). As Warschauer (1996) points out, motivation and engagement also play a crucial role in the effective use of technology in language learning. In addition, learners' prior experience with technology enhances the learning process as they become familiar with the platforms and methodologies involved (Leakey, 2011). Individual differences, however, can also lead to distractions with digital technology, requiring teacher management to maintain focus. (Rosen, Carrier, & Cheever, 2013). Identifying differences of this kind, through analyses and observations, can aid educational institutions in significantly improving the language learning experience.

A thorough assessment of studies on AWE tools from 2000 to 2020 was undertaken by Nunes et al. (2022), displaying a noteworthy dearth of studies investigating the the long-term effects of AWE tools. Despite the fact that these tools clearly offer advantageous effects to both teachers and students, it is important to acknowledge the significance of examining their long-term effects on students' writing skills. If the feedback mechanisms implemented by AWF tools fail to assist students in retaining their writing abilities and transferring the corrections made into their writing practice, then the influence of these tools on students' actual writing may only be superficial, comparable to the automatic error correction functions present in common word processing software. Therefore, additional investigation must be conducted to confirm whether AWF tools help or weaken students' writing skills. Through a meta-analysis of 26 studies carried out between 2010 and 2022, Zhai and Ma (2023) aimed to shed light on the influences of automated writing evaluation on writing quality. Collectively, the results showed that AWE tools have a typically beneficial impact on students' writing abilities. Particularly, compared to students who received feedback using conventional techniques or none at all, 86% of intervention group participants displayed stronger writing skills. AWE tools were additionally found to be more advantageous for EFL and ESL learners than native English-speaking students, according to the study. The analysis of the impact of AWE tools on different writing genres revealed that they significantly improved argumentative writing, while having a less pronounced effect on academic writing and diverse mixed genres. While the researchers acknowledged the beneficial effects of AWE tools, they also highlighted that the length of the intervention and the use of different feedback techniques and AWF platforms to improve writing quality had only a limited impact. In order to narrow the gap between genre writing and disciplinary conventions, Cotos' (2014) work focuses on utilising specialised corpora and AWE to improve L2 research writing teaching. By employing explicit language analysis, customised corpora, and AWE technology to promote genre-specific knowledge acquisition and metacognitive awareness, this initiative aims to overcome educational issues regarding writing feedback, particularly with varying genres. The study places a strong emphasis on integrating genre-oriented AWE to deliver genuinely individualized feedback and support L2 students in efficiently internalizing and utilizing genre conventions. Finally, the research suggests that by providing a strategic method for teaching

discipline-specific genre knowledge and increasing L2 writers' academic discourse abilities, L2 writing instruction could be advanced. The viability of genre-oriented automated writing evaluation in postsecondary education is investigated by Burstein et al. (2016). The study investigates the potential for AWE to address writing requirements and faculty perceptions of student writing proficiency across diverse genres. The results highlight the need for AWE systems to accommodate a larger range of writing tasks, such as research proposals, in order to better align with both academic and professional expectations. This genre-based methodology encourages the creation of a thorough taxonomy for assessing writing quality across genres. In order to facilitate more comprehensive writing teaching and better prepare students for complex writing tasks, the study suggests the use of modified AWE tools that incorporate genre-specific conventions, coherence, and source integration. While useful in improving AWE ability to assess genres, there is potential for further improvement in reliability and consistency of automated writing feedback through the use of specialized corpora dedicated to varying student levels, subjects, and purposes.

As with any emerging technology that is constantly evolving and being improved, a key consideration with both AI and AWE is that, as Grimes and Warschauer (2010) point out in their study, such technologies should be viewed as 'fallible' and should not be blindly trusted, by teachers and administrators, but also students and others who may benefit from them. As noted above, it is vital that further research is conducted on these ever-improving tools, bearing in mind that the drawbacks and limitations identified in earlier studies may have been addressed and improved since the time of publication.

2.4.1. Exploring the Benefits of AWE

Automated writing evaluation tools, alternatively referred to as automated essay evaluation, automated essay scoring, or automated feedback tools (Hockly, 2019; Warschauer & Ware, 2006), were initially devised to provide summative scores for written essays in high-stakes examinations. Over time, however, these systems have undergone substantial modification and adaptation to facilitate their integration into classroom teaching and assessment practices. As well as providing objective scores, AWE systems have been refined to provide automated feedback to students. The development and implementation of AWE systems has been extensively investigated by researchers such as Shermis et al. (2016), and Stevenson andPhakiti (2014), contributing to their continued improvement. As a result, AWE tools continue to evolve, periodically introducing new features and improvements to maintain their role in supporting writing pedagogy and assessment. Eventually, a distinction was made by scholars between automated scoring and automated feedback, which has been addressed by researchers such as Burstein et al. (2020), Shermis et al. (2016), Stevenson and Phakiti (2019), Ware (2014), and Xi (2010). Research has demonstrated that AWE systems can provide reliable and valid measures of writing ability (e.g., Shermis, 2014). Zhai and Ma's (2022)

analysis confirms the notion that AWE tools can offer a reliable new tool for teachers and students, while also highlighting their ability to assess and improve different genres of writing.

AWE been extensively studied in testing contexts, with research focusing on the validity and reliability of scoring systems, culminating in optimistic results (Attali, 2013; Shermis & Hamner, 2013; Shermis et al., 2002). Several studies have demonstrated that AWE systems can provide reliable and valid measures of writing abilities (e.g., Shermis, 2014), leading to their use as second raters in scoring the writing components of high-stakes tests such as the GRE, TOEFL, and placement tests (Williamson et al., 2012). Perelman (2014) criticized automated essay scoring (AES) systems for an apparent large bias towards word count, referring to them as merely more than "state-of-the-art" word counters (abstract). Although initial concerns were raised about the presence of "bugs" or exploitable functions in AWE tools (Powers et al., 2002; Perelman, 2014), recent studies (Wilson & Rodrigues, 2020; Kumar & Boulanger, 2021) suggest that developers have likely addressed and resolved this issue thanks to users' and researchers' feedback. Wilson and Rodrigues (2020) specifically highlight that human raters' scores were also influenced by word count, as students falling below a specific word limit incurred point deductions. Research has also explored the relationship between AWE and human raters, showing significant negative longitudinal performance changes as opposed to those of human raters (Lamprianou et al., 2020). Studies on peer feedback and AWE suggest that when used together, AWE tools have the potential to improve writing quality and foster improvement (e.g., Warschauer & Ware, 2006; Huang & Renandya, 2020). AWE tools have also been considered efficient and time-saving tools, providing feedback comparable to that of teachers (Wilson et al., 2021). Additionally, AWE tools with scaffolding practices have shown promise in improving students' learning motivation and autonomy while easing teachers' workload (Cotos, 2011). Students have reported positive outcomes with AWE feedback, finding it helpful in identifying language gaps and errors and enhancing cross-linguistic development (Mohsen, 2022). Furthermore, the customized feedback provided by AWE functions as a mode of mediation, assisting in students' writing processes and facilitating the cultivation of proficient cognitive writing frameworks (Jiang et al., 2020). Recent investigations have also explored the impact of automated feedback onEFL students' writing quality. Cheng (2017) found that online automated feedback positively influenced the reflective journals of university-level EFL students, leading to improved writing quality. Similarly, Chew et al. (2020) conducted a comparative study on ESL students' summary writing, showing that the use of Summary Writing-Pal (SW-PAL), an automated tool, led to enhanced summary writing quality compared to traditional methods. In a classroom setting, Hassanzadeh and Fotoohnejad (2021) examined the efficacy of Criterion, an automated writing evaluation tool, and found that its feedback resulted in significant improvements in the overall quality of writing for EFL students compared to teacher-generated feedback.

These findings demonstrate the positive effects of automated feedback tools in enhancing the writing proficiency and development of EFL students. However, future research should address

concerns raised by Perelman (2014) and explore the potential limitations and nuances of AWE scoring systems to ensure their continued effectiveness in writing assessment. While the findings on the credibility of AWE could serve as a testimony of their approval, nonetheless it would be optimal for educators and policymakers to test and evaluate the tool themselves before committing to implementation.

2.4.2. Ethical and Practical Considerations on AWE

Automated writing assessment and evaluation systems have met with mixed opinions and criticism in the literature, in spite of their evident benefits. The reliability of automated scoring, a main component of AWE, has received extensive coverage in the literature (Shermis et al., 2002; Lee et al., 2008; Kellogg et al., 2010; Attali, 2013; Shermis & Hamner, 2013). While AWE tools offer potential solutions to existing challenges in writing instruction and feedback, they also present their own set of difficulties, including the potential for misuse and overuse. Chen and Cheng (2008) conducted a study that revealed negative evaluations of the My Access tool that raised pertinent questions concerning the efficiency and suitability of automated feedback tools in educational contexts. Their findings highlight the importance of adequate teacher training and preparation for the implementation of such tools, a point also emphasised by Grimes and Warschauer (2010), as well as the crucial role of student attitudes (El Ebyary & Windeatt, 2010; Wilson et al., 2021) and acceptance (Davis, 1985; Zhai & Ma, 2022). However, with the advancements in AI and NLP technology and the emergence of newer AWE methods, it is gradually becoming clear that these technologies are not intended to replace teachers, but rather to complement their roles (Wilson & Czik, 2016; Celik et al., 2022; Link et al., 2022). However, the integration of any tool into education should not be pushed solely based on its advantages in one aspect, while disregarding potential drawbacks in another. It is without doubt that to conduct thorough examinations into the long-term ramifications of AWE utilization is imperative. The emphasis of AWE implementation should be directed towards examining the potential for positive transfer of writing skills. The results of the study of Nunes et al. (2022) support the notion that AWE tools are a reliable source for both teachers and students, demonstrating their effectiveness in assessing and improving various styles of writing. Yet, the researchers acknowledged the existence of findings that were either unfavourable or inconclusive (Wilson & Czik, 2016; Cohen et al., 2018; Wilson & Roscoe, 2020), highlighting the need for further research and enhancements to the application of AWF techniques to optimize their potential advantages. Indubitably, further investigation into these dynamically evolving and enhanced tools is imperative to evaluate their current state comprehensively.

While Li et al.'s (2014) research revealed a generally positive predisposition towards the Criterion software, students voiced some particular difficulties with the tool. These included, in particular, a striking similarity in the formulaic score descriptions provided alongside the assigned scores, regardless of whether the assessment was 4 or 5 out of 5. In addition, student feedback

indicated that the scores failed to improve in any significant way after correcting reported errors. In a similar vein, Grimes and Warschauer's (2010) three-year study of My Access in eight US middle schools explored the attitudes of teachers and students towards the use of the automated writing tool. Despite the fact that many teachers expressed doubts about the reliability of automated scoring and that teachers' use of the software was limited by their preference for traditional writing methods, My Access did offer certain benefits, such as efficiency, increased writing output, motivation and revision. However, the study identified limitations, including the software's inability to assess organisation and content, and occasional inaccuracies in scoring and corrections. One teacher reported that correcting students' major errors after My Access had marked them resulted in even lower grades, which potentially discouraged students, while another student's score dropped due to the accidental inclusion of a single word, which undermined confidence in the software. In addition, as previously observed in a study of Criterion by Grimes and Warschauer (2010), students who were more digitally literate found techniques to cheat the software, e.g. by repeating words or making essays excessively verbose to inflate the word count. Regardless, as teachers gained experience with My Access, their confidence in the tool also grew, resulting in a significant improvement in the number of revisions made, from 12% to 53%, as reported by Grimes & Warschauer (2010). Although most teachers did not rely on My Access for marking, they appreciated its capacity to provide feedback on the mechanical aspects of students' essays, freeing up time to focus on more complex writing issues. However, in their study investigating the appropriateness of AWE holistic scores for classroom-based formative assessment in a university-level English writing course, Li et al. (2014) revealed that AWE holistic scores, when properly implemented and used for their intended purpose, showed high correlations with human assessment in a variety of scenarios. Correlations between AWE scores, instructor ratings, and ESL instructors' and students' perspectives are assessed through the Criterion tool. The results show a moderate but positive correlation, suggesting that AWE scores may be useful in identifying students' writing needs.

In the realm of English as a second language (ESL) education, the utilization of AWE software which utilizes holistic scores has become increasingly prevalent in recent years. A study undertaken by Li et al. (2014) examined the efficacy of such tools within the context of ESL classrooms. Specifically, the study sought to explore the correlation between the scores provided by Criterion, and the grades and ratings assigned by instructors, in addition to assessing the perceptions of both instructors and students regarding the tool. The study's findings revealed a moderate correlation between the scores assigned by teachers and those provided by Criterion. An analysis of the scores also demonstrated that Criterion was significantly correlated with teachers' scores in all categories with the exception of material. However, despite this correlation, instructors participating in the study expressed minimal trust in the scores provided by Criterion. It is worth noting that this trust was galo limited to low scores. In contrast, the students' perspectives regarding Criterion were largely positive, with moderate to high levels of trust in the scores assigned by the tool. Nevertheless, despite these criticisms, students found the scores assigned by Criterion to be a strong source of motivation,

leading to a sense of accomplishment when receiving higher scores. Instructors, on the other hand, utilized the scores provided by the tool as a warning system for students with weaker writing skills, paying increased attention to those who received low scores from Criterion. In addition, they implemented a threshold system that necessitated students to attain scores above a certain level on Criterion to be able to submit their essays.

In spite of the robust body of evidence supporting the effectiveness of automated essay evaluation (AEE) and AWE systems, empirical studies have occasionally yielded insignificant or null outcomes. Notably, the study conducted by Wilson and Czik (2016) did not reveal statistically significant variations in the final-draft writing quality between the two groups, despite the discernible inclination of students benefiting from automated tools to exhibit heightened levels of writing improvement when compared to their counterparts exposed solely to teacher feedback. This line of inquiry is further bolstered by the study of Wilson and Roscoe (2020), which employed multifaceted criteria to gauge the efficacy of the Project Essay Grade (PEG) in the realm of writing assessment, ultimately reporting comparable findings of minimal import.

Vojak et al. (2011: 108) assert that the complex nature of writing, characterized by its “socially situated”, “functionally and formally diverse”, and “increasingly multimodal” attributes, necessitates human interaction following a thorough evaluation involving a range of AWE tools, including Criterion, My Access, Essayrater, and PEG. Additionally, their research showed evidence of formulaic techniques, feedback that lacked specific details, incorrect detection of errors, a strong emphasis on writing mechanics such assyntax and punctuation, and a tendency to value length over meaningful content, which was also criticized by Perelman (2014). Vojak et al. (2011) ultimately determine that the challenges might be overcome with technological advancements, but only if students are encouraged to explore said attributes throughout the process.

In Lai's (2010) study, students expressed frustration that the feedback provided by My Access was too vague and sometimes repetitive, making it impossible for them to effectively revise and improve their writing. Yang (2010) also investigated My Access from the perspective of students' opinions and attitudes. Although opinions were generally positive, the tool was similarly criticised for being repetitive and having issues with accuracy. Other criticisms of My Access were also made by Chen and Cheng (2008), who cited the lack of original input, absence of specific suggestions and perceived unfair grading. Despite these criticisms, other research (Yeh et al., 2007) has found that the use of My Access has positive effects and benefits. Importantly, more recent research has supported the efficacy of the tool, highlighting its effectiveness both in terms of student engagement (Chou et al., 2016) and as a means of feedback (Mohsen & Abdulaziz, 2019) in helping students. The improvements to My Access since Lai's original study in 2010 may account for these more recent findings.

In a study conducted by Ware (2014), the impact of various feedback types on the writing output of 82 adolescent students in a secondary English class was examined. The students were divided into three groups, each receiving different forms of feedback: automated feedback from the writing assessment software Criterion, electronic feedback from the teacher, or peer feedback on paper drafts. The findings indicate that all students exhibited improvements in writing length and holistic scores after the 6-week intervention, irrespective of the feedback type received. Notably, peer and electronically delivered teacher evaluations significantly enhanced students' performance in the open-ended answer genre. Conversely, Criterion proved less effective than peer and instructor evaluations in improving the length, quality, and mechanics of writing. Students who received automated feedback via Criterion did not perform as well in their writing assignments compared to those who received feedback from peers or teachers. These results underscore the importance of carefully considering different feedback techniques in secondary English classrooms and advocate for further research to comprehensively understand their benefits across various educational settings. Warschauer and Ware (2006) emphasize a critical consideration: several studies on Automated Writing Evaluation (AWE) tools have been funded by the tools' publishers. To ensure transparency and impartiality, it is recommended that future research in this domain seek independent funding and oversight.

Hockly (2019) explored the suitability and validity of AWE tools for their intended purposes, such as summative assessments and formative writing assistance, and highlights that although such technologies will likely continue in their existence and use, their limitations and effectiveness might negatively impact use. To address such issues, Kane (2013) recommends an argument-based validation approach for performance-based evaluations, including those associated with AWE tools. This approach involves creating a network of arguments and inferences to thoroughly evaluate the accuracy of the assessment. Chapelle et al. (2015) have proposed a validation framework for the utilization of AWE with second language (L2) writers. The framework comprises seven domains: evaluation, generalization, explanation, extrapolation, utilization, ramification, and domain definition. The effectiveness of AWE can be influenced by various factors, such as language proficiency, pedagogical approaches, and attitudes towards its use (Zhai & Ma, 2022). Empirical studies have shown that learners with advanced linguistic analysis skills tend to benefit more from regular written corrective feedback (CF) targeted at specific linguistic features (Sheen, 2007). To obtain a more comprehensive evaluation of writing, it has often been suggested that AWE should be merged with other assessment systems. Further research, inspired by the findings of Kane (2013) and Chapelle et al. (2015), will be essential to promote the integration of AWE in the field of ELT. Because both AWE and the feedback generated by such tools originate from computational sources, their ability to assess and evaluate texts is inevitably limited by the programming provided by their developers. Although many educators and students find AWE user-friendly, valuable, and accurate (e.g., Wilson & Roscoe 2020; Han & Sarı, 2022), a body of literature also highlights distinct shortcomings of AWE tools compared to human instructors in assessing higher-level skills such as

organisation, style, content, and ideation (Dikli & Bleyle, 2014; Wilson & Czik, 2016). Furthermore, AWE tools have been subject to criticism for their limited ability to provide individualised and targeted feedback. Li et al. (2014) express concern that the source data and feedback from AWE tools may have lacked originality in their study. They argue that the tools analyzed and proposed changes that guided students' work to fit a fixed pattern, which could potentially be detrimental to students' creativity and independent thinking. Such shaping of student writing can be indeed a dangerous example for students who seek feedback from a mechanism designed to be accurate, fair and reliable. There have been reports that automatic feedback systems have difficulty meeting the personal needs and learning preferences of students, and it is noted that visual and auditory feedback may be more appropriate. Furthermore, personalised feedback provided by human teachers is crucial in assisting students who face writing difficulties (Koenka & Anderman, 2019). Considering the integration of automated feedback systems, the cost and availability of such systems should also be taken into account. Additionally, the objectivity of these systems depends on the algorithms and data sets they are built upon. Thus, biases within these algorithms and datasets may result in biased feedback, particularly affecting non-native English speakers, or those whose writing patterns are not aligned with that of the dataset. While automated systems excel in addressing grammatical and syntactical aspects of writing, they may fall short in providing guidance on more intricate matters such as argumentation, creativity, and individual voice.

The notion of “gaming” AWE tools refers to the practice of tricking the AI system intending to elicit higher scores than those warranted by the content's quality or deservingness. Gaming methods used and potentially could be used by students are examined thoroughly in the literature with a variety of tools being tested for vulnerability. (Powers et al., 2002; Perelman 2014; Klebanov et al., 2017;). The study by Powers et al. (2002) investigated the efficiency of e-rater, an automated essay grading system, by analyzing its vulnerability to manipulation. Participants were instructed to write articles that attempted to game the system, indicating an uneven distribution of success. With findings that display that scoring through e-rater could be tricked into giving high scores, the study encouraged AWE collaboration with human assessors and stressed the need for care when using AES systems exclusively for high-stakes exams. The study also emphasized the differences in the evaluations of several essay characteristics by e-raters and human readers, highlighting the necessity of human intervention to correct any potential errors. The researchers recommend a supplementary strategy in which computerized systems concentrate on specific markers of writing quality to assist human raters. Higgins and Heilman (2014) also discuss the potential risk of AWE tools perpetuating counterproductive writing strategies in students, including the incorporation of manipulation methods such as employing superfluous language or merging academic and non-academic terms within their essays.

Concerns have been raised regarding the potential difficulty in utilizing AWE tools for individuals with limited experience in computers and AI tools, as opposed to those with more

advanced knowledge in these areas (Zhang & Hyland, 2023). Alternatively, individuals with experience with how such tools work can purposefully alter their writing to conform to the standards that AWE systems consider indicative of a well-written essay, therefore receiving an undeserved positive evaluation. Students who are more skilled with computers can potentially manipulate AWE systems to their benefit as well. At the same time, their familiarity with computer tools such as keyboard shortcuts and writing assistants such as Grammarly, Quillbot and ChatGPT could serve as advantages. As such, it is essential that all students receive instruction on these techniques in order to ensure fairness and equitable opportunities. Furthermore, present concerns about cheating AWE systems highlight the need for comprehensive instruction for students and effective instructor supervision, emphasizing the critical role that administrators and teachers play in resolving these problems. While traditional AWE may have historically focused on surface-level language features like essay length and sentence structure, recent advancements in AWE technology have successfully integrated more intricate linguistic elements, including in-depth language nuances and world knowledge (Deane, 2013). In the study by Lai (2010), concerns were raised about the adequacy of feedback provided by AWE tools, particularly in terms of consistency, unification, and relevance to students' needs. Drawing a parallel, Rudolph et al. (2023) aptly reflect on the historical concern surrounding handheld calculators and the potential decline in numeracy skills. Similarly, the apprehensions surrounding AWE tools should be balanced by recognizing the iterative nature of their development. While limitations and challenges exist, a proactive approach involving proper training, user input, and continuous refinement offers the promise of unlocking substantial benefits through AWE tools. It is critical to acknowledge that there is not yet a feedback tool that can provide the same level of detail as a teacher can, such as tracking the gradual development of the learner, predisposition, experience, strengths and weaknesses, and the recognition of recurring and potential errors that are inherent in the feedback a teacher provides.

2.4.3. Learner Engagement with AI Assistance

In the recent study by Fu and Liu (2022), the effectiveness of AWE feedback provided to EFL learners is examined as well as their engagement. While the effectiveness of AWE in improving writing correctness is acknowledged, this research focuses on the area of behavioural, cognitive and affective learner engagement with AWE feedback, which is comparatively understudied. The goal of the study is to gain a thorough understanding of the aspects of learner engagement as well as how it influences learning outcomes, particularly in AWE feedback. The results highlight the manner in which these engagement components interact intricately and how they enhance AWE's effectiveness. The researchers acknowledge that learners' cognitive and emotional responses seem to possess an impact on their degree of involvement with feedback. This interconnection is reported to contribute to the effectiveness of AWE in improving learning outcomes. The study encourages rational methodological strategies, such as mixed methods, to accurately measure the progression of learner engagement. In Hostetter et al. (2023), faculty members have expressed concerns about the

application of AI tools and their possible effects on assessment methodologies. The study emphasizes that understanding AI technologies and improving communication over the goals of writing tasks are crucial components that have emerged. The study emphasizes that while AI shows promise, encouraging informed discussions and offering guidance are essential for ensuring effective learning experiences. Notably, the study reveals gaps with probable inconsistencies in the quality of text produced by AI and the rapid growth of AI technologies. Previous investigations have delved into the impact of employing AI-based writing assistants on the English writing proficiencies of EFL students (Alharbi, 2023) and the attitudes of students towards artificial intelligence (Keleş & Aydın, 2021). Keleş and Aydın (2021) observed that university students generally held a negative viewpoint toward artificial intelligence, associating them with war, laziness and lacking in judgement. While such negative sentiments toward AI may also apply to today's LLMs, it is important to also consider that at the time of the study, a majority of the participants of the study thought of Apple's Siri or Sophia, developed by Hanson Robotics. Another focal point of inquiry has been computer anxiety (Matsumura & Hann, 2004) or computer apprehension (Yaghi & Abu-Saba, 1998), which stems from perceived inadequacy in computer literacy and past adverse experiences. Despite the prevalent high levels of computer literacy among contemporary students, the introduction of novel technologies, user interfaces, and software that demand dedicated effort and time for proficiency acquisition can potentially evoke computer anxiety (Russell, 1995; Mac Callum & Jeffrey, 2014). Correspondingly, Lai (2010) identified that participants in their study encountered complexity and confusion while navigating the multifaceted functions of My Access, leading to a sense of unease. When educational software is difficult to navigate and use, it impedes the achievement of tasks and tends to be disliked, regardless of the speed and educational practicality it offers (Faghih et al., 2014). To ensure the effective utilization of new tools and software without the encumbrance of anxiety and perplexity, it would be beneficial for educators and administrators to furnish students with guidance and training in the effective use of automated feedback tools and educational supplementary technologies. Furthermore, a prudent recommendation for the developers of emerging technologies is to undertake usability tests with the intended user base before public release, thereby scrutinizing user experiences comprehensively. Hence, teachers' adeptness in navigating new technologies and tools that students may encounter or could benefit from throughout their academic journey becomes an essential consideration. While a student's cultural background can indeed influence various dimensions of their learning experience, it could also be a determinant of the suitability of a particular feedback mechanism (Nelson & Murphy, 1993; Carson & Nelson, 1994; Nelson & Carson, 1998). Even in situations where automated feedback falls short of student and instructor expectations or has proven unable to outperform other kinds of feedback, it remains a promising supplementary avenue that can offer advantages to those who can effectively harness its potential (Lai, 2010).

2.4.4. Acceptance of New Technologies

A review of the development of technology throughout history reveals that innovations have often been met with criticism due to worries that they could impede independent and critical thought. Examples of this include search engines, and online encyclopedias such as Wikis (Matthew et al., 2009; Meishar-Tal, 2015). The effects of perceived usefulness (Venkatesh et al., 2012) on automated feedback and the technology acceptance of students were examined by Zhai and Ma (2022). Several factors that may positively or negatively affect technology acceptance were found to be teachers' preparedness, ability and impression, as well as technology literacy, availability of Technology and students' backgrounds. These findings as well as a majority of the studies on the Unified Theory of Acceptance and Use of Technology (UTAUT) have shown that the perceptions of individuals on the usefulness, effectiveness and value of new technology affects their acceptance of it suggest that the first impressions created throughout the initial introduction of AWE in classroom settings should be comprehensible, interesting and free of areas of confusion. Innovator's dilemma is defined as the tendency to accept that which is new (Clayton Christensen et al., 2011). A similar concept is investigated by researchers through technology acceptance of new technologies (Davis, 1989; Legris et al., 2003; Strzelecki, 2023). The use and acceptability of AWE by Chinese college students in the setting of CALL were examined by Li et al. (2019). In order to examine the complex relationships between learners' willingness to use AWE and variables of the technology acceptance model (TAM), such as perceived usefulness, attitude, perceived ease of use, computer self-efficacy, and computer anxiety, the authors incorporate external factors computer self-efficacy and computer anxiety into the TAM. The findings imply that perceived usefulness, attitude, and computer self-efficacy were the main factors influencing learners' desire to use AWE, with perceived ease of use having a relatively less but still substantial impact. It was highlighted in the findings that although computer self-efficacy and computer anxiety were found to affect usability, they do not directly influence the opinion of usefulness.

Considering that some learners may face difficulties in adopting new technology due to preconceived notions, the concept of technology acceptance should be considered by teachers and researchers before proceeding with the testing or implementation of new tools or methods that involve technological innovations. Such considerations are crucial for the effective incorporation of genAI and AWE tools into educational settings.

2.5. Research Gaps in AWE and GenAI

Fu and Liu (2022) emphasize the need for greater clarity of learner engagement with AWE feedback by pointing out gaps in the literature. They draw attention to the understudied emotive and psychological aspects of involvement as well as the dearth of mixed-method research. To fill these

gaps and further develop the comprehension of learners' engagement with AWE feedback, the authors urge further study employing scientific instruments and larger-scale inquiries.

Han and Sarı (2022) also highlight a number of unexplored topics in the field of AWE. Despite being widely used in K–12 and higher education settings, AWE systems' ability to improve writing development and user perceptions was reported to be still unknown, which can be partly related to the divergent research designs used. While certain investigations utilize within-group designs to contrast automated feedback against no feedback or teacher-generated feedback, others use within-group designs to compare student performance across submissions. A dearth of studies were reported to examine the beneficial effects of combining automated and instructor input. The researchers also reported that further empirical investigation is needed to determine the usefulness and validity of AES in low-stakes writing examinations and classroom-based assignments, particularly in non-English environments such as Turkish schools. To determine AWE applicability and accuracy in various educational situations, these research gaps call for thorough investigation. Within the present review of literature, the exploration of prompt engineering in conjunction with ChatGPT appears constrained, particularly within educational contexts. Nevertheless, despite the limited research available, the potential of prompt engineering in simplifying and optimizing tasks within education for learners, teachers, and researchers is apparent. Further research is required, and likely will be present in the future years, to fully investigate and harness the benefits of prompt engineering with ChatGPT in educational settings.

Ranalli and Yamashita (2022) suggest that further investigation into AWE should encompass developments in error-correction, particularly with collocation issues. In the context of specifically designed tools, Saricaoglu's (2019) study investigates the impact of automated formative feedback on the improvement of ESL learners' written causal explanations, through an investigation of a specifically designed tool. The research shows that while automated feedback can produce significant improvements within individual essays, this may not necessarily transfer to long-term development across multiple tests. As referred to by Saricaoglu, a significant study gap exists in understanding how long-term exposure to AWE tools affects L2 students' autonomous error recognition and correction abilities (Ranalli & Yamashita, 2022).

Future research in the area of AWE, according to Wang et al. (2013), should continue to consider the individual differences among students, particularly how they react to feedback, how interactive they perceive it, and how much autonomy they feel they have over their education. In order to tailor AWE tools that enhance the effectiveness of automated feedback and the overall learning process in EFL writing classes, it may be essential to acknowledge such differences. The limited amount of research focusing on the capacity of AWE tools to evaluate different writing genres raises concerns about their usability and underscores the need for comprehensive investigation (Cotos, 2014; Ware, 2014; Zhai & Ma, 2022).

Another gap in the literature was the absence of research examining Paperpal. Despite being recognized by educators and researchers as a valuable resource for academic writing and corrective feedback, the effectiveness of the Paperpal application remains unexplored in existing studies, especially in comparison with other similar web-based writing assistance applications such as Grammarly (e.g., Calma et al., 2022; Thi et al., 2023) and Quillbot. Despite the recognized efficacy of AWE tools in enhancing linguistic accuracy within EFL domains, the literature presents a gap in the exploration of the influence of such technologies on student motivation and anxiety—critical psychological dimensions that affect successful learning outcomes (Fu & Liu, 2022; Alharbi, 2023).

While investigations, such as those by Li et al. (2019), aim to untangle the interplay between technological acceptance and variables like computer self-efficacy and computer anxiety, comprehensive analyses that delineate how AWE feedback specifically affects learner motivation and anxiety remain scarce. This oversight is notable given the emerging discourse on the potential of AI to revolutionize educational practices and its challenges (Hostetter et al., 2023) and the criticality of addressing both the cognitive and affective facets of learner engagement to harness the full pedagogical potential of AWE technologies (Keleş & Aydın, 2021; Mac Callum & Jeffrey, 2014).

The academic community's call for enriched methodologies such as mixed-method approaches to accurately chart the progression of learner engagement with AWE feedback (Fu & Liu, 2022), alongside an improved understanding of the underlying psychological constructs, notably motivation and anxiety, is imperative for advancing the educational efficacy of AWE systems. This observation is particularly intriguing given that these facets have been scrutinized in the context of research pertaining to writing performance (Morphy & Graham, 2012; Öğüt, 2022; Shang, 2022) and efficacy in relation to various AWE tools (Foltz et al., 1999; Burstein et al., 2004; Grimes & Warschauer, 2010; Li et al., 2014).

2.6. Conclusion

Recent research explored the integration of AI tools across scholarly domains, with ChatGPT's applications and impacts under scrutiny (Alshami et al., 2023; Wang et al., 2023). Ethical concerns were raised regarding AI-generated content and potential misuse (Hostetter et al., 2023). Positive student responses in education were noted, while ChatGPT's potential in search engines was envisioned (Aktay et al., 2023; Alafnan et al., 2023). The assessment of ChatGPT's language complexity was examined, and its assistance in literature reviews was explored (Wang et al., 2023). The balance between ChatGPT and human writers was highlighted (Yan, 2023; Zhou et al., 2023). AWE tools' benefits, including timely feedback, were acknowledged, while challenges in content aspects were identified (Stevenson & Phakiti, 2019; Wang et al., 2013; Ngo et al., 2022). Motivation and anxiety were addressed using AI tools (Ali et al., 2023; Li, 2022). Over-reliance and cheating

concerns persisted (Bawarith et al., 2017; Kasneci et al., 2023). The educational impact of AI integration was nuanced, addressing authenticity, ethics, and bias (Göçen & Aydemir, 2020; Graham, 2022; Aydın & Karaarslan, 2023; Fallon, 2023; Kumar, 2023). Critical thinking amid ChatGPT's inaccuracies was emphasized (Biswas, 2023; McKenna et al, 2023). AWE's evolution from scoring to feedback was outlined (Shermis, 2014). Balancing AI feedback and human guidance posed a challenge (Fu & Liu, 2022). Overall, ethical, educational, and technological considerations framed the discourse around AI's integration.

Current literature highlights the crucial role of motivation in language learning, particularly noting a shift towards extrinsic motivational factors. Studies conducted by Wong (2014) and Yalçın (2018) show that career aspirations and participation in extra-curricular activities serve as significant motivational factors for students in secondary school and preparatory school. This pattern is similarly observed among Indonesian students (Astriningsih & Mbato, 2019) and Vietnamese university students (Ngo, 2015), where extrinsic motivations predominate, although intrinsic factors also positively influence overall motivation. Research on demotivation shows that anxiety and environmental conditions are key contributors, with male students showing higher levels of demotivation than their female counterparts (Acat & Demiral, 2002). The relationship between gender, motivation and anxiety is further supported by the findings of Çetin et al. (2020), who indicate that female students report higher levels of both. Furthermore, Yıkılmaz (2020) highlights that despite the dominance of extrinsic motivation, intrinsic factors such as personal interests in communication and travel contribute significantly to students' motivation. The integration of AI tools in language learning has shown encouraging results. Research by Nazari et al. (2021) suggests that applications such as Grammarly improve student engagement and self-efficacy. In addition, Ali et al. (2023) show that ChatGPT can increase motivation for reading and writing tasks, although students express ambivalence about its effectiveness in developing speaking and listening skills. However, Buriak et al. (2023) caution that AI-generated content requires careful scrutiny and should not be relied upon alone. This highlights the ongoing need to critically evaluate the effectiveness and limitations of AI tools in educational settings.

Extensive research on ChatGPT emphasizes its potential within education and beyond, identifying critical gaps and concerns. In light of its substantial benefits, such as increased productivity and enhanced accessibility in education, the limitations and ethical implications of ChatGPT—including potential misuse and concerns regarding accuracy—underscore the necessity for careful integration. The existing literature calls for further empirical research to quantify the impact of ChatGPT, establish best practices, and address ethical considerations, particularly regarding academic integrity. While AI tools like ChatGPT present transformative opportunities, their implementation necessitates thoughtful management to optimize benefits while mitigating associated risks.

AWE software has evolved significantly over the past decade, improving the efficiency and effectiveness of writing feedback in educational settings. While early tools such as PEG laid the groundwork, modern AWE systems such as Grammarly and Paperpal use advanced technologies and machine learning methods to provide immediate, detailed feedback. These systems can reduce teacher workload and offer timely corrections, contributing to improved writing skills and student motivation. However, AWE tools have their limitations. They may not fully address the nuanced aspects of writing, such as argumentation and content structure, and their effectiveness can vary depending on the user's ability and length of use. Studies suggest that while some tools are highly effective, others show mixed results or are less reliable. In addition, the integration of AWE systems requires careful consideration of their limitations and the educational context in which they are used. Overall, AWE systems represent a valuable resource for improving writing instruction, but they must be implemented thoughtfully, with attention to their strengths and weaknesses. Continued research and refinement are needed to optimise their benefits and ensure that they complement, rather than replace, traditional teaching practices.

This chapter discussed the integration of AWE and LLM tools in education and research, emphasizing their potential benefits and challenges. It examined issues like authenticity, ethics, and the need for a balance between automated feedback and human guidance. Additionally, the chapter delved into the impact of said tools on L2 learners' anxiety and motivation, highlighting the importance of considering these factors in the implementation of AI tools. It is suggested that future research devote particular attention to the long-term effects of AWE tools on students' abilities to write independently. Investigations on the effects on engagement, anxiety, and motivation are exceptionally vital and should be thoroughly explored. Along with a careful analysis of individual differences, it is also suggested to recognize the crucial importance of instructional methodologies and students' digital literacy skills.

CHAPTER THREE

3. METHODOLOGY

3.1. Introduction

This chapter provides an examination of the methodology utilized in the study. It encompasses the research design, approach, sampling methods and tools employed for data collection, accompanied by insights into the decision-making process that led to the selection of these specific methods. The research questions guiding the study are introduced, and each phase of the research is meticulously expounded upon to ensure transparency and serve as a guide for future studies. The chapter begins with an outline of the chosen research design and elucidates the rationale for its selection. Subsequently, it delves into the setting, participants, and instruments employed in the study. Finally, it provides a comprehensive overview of the gathered data and the methodologies employed for its acquisition. This detailed description is crafted to assist future researchers in replicating and effectively implementing these methods.

3.2. Research Design

This case study employed a quasi-experimental mixed methods approach to thoroughly investigate the impact of various feedback methods on writing anxiety and motivation among English students. This methodology was selected to capture an integrated understanding of students' experiences with This study employed a quasi-experimental mixed methods approach to thoroughly investigate the impact of various feedback methods on writing anxiety and motivation among English students. As Creswell and Creswell (2017) articulate, mixed-methods research, which integrates qualitative and quantitative data, provides a comprehensive understanding of a research subject. They emphasize that “to include only quantitative and qualitative methods falls short of the major approaches being used today in the social and human sciences”. This methodology was selected to capture an integrated understanding of students' experiences with Automated Writing Evaluation (AWE) tools, specifically PaperPal and ChatGPT, and their subsequent effects on writing anxiety and motivation.

The quasi-experimental design of the study was chosen due to the pre-existing nature of the sample, which precluded random assignment to control and experimental groups (Campbell & Stanley, 2015). To ensure that the groups did not significantly differ in academic achievement, t-tests

were conducted on their preparatory exemption exams (See Table 2). The combination of semi-structured interviews and surveys provided both numerical insights and contextual narratives, thereby enhancing the depth of understanding of the phenomena under investigation. This comparative design examined the differential impact of feedback methods, where one group received feedback from teachers and the other received AWE feedback. The findings from this study contribute significantly to the understanding of how different feedback mechanisms influence students' emotional and motivational responses within the context of English language education, Automated Writing Evaluation (AWE) tools, specifically PaperPal and ChatGPT, and their subsequent effects on writing anxiety and motivation.

The quasi-experimental design of the study was chosen due to the pre-existing nature of the sample, which precluded random assignment to control and experimental groups (Campbell & Stanley, 2015). To ensure that the groups did not significantly differ in academic achievement, t-tests were conducted on their preparatory exemption exams (See Table 2). The combination of semi-structured interviews and surveys provided both numerical insights and contextual narratives, thereby enhancing the depth of understanding of the phenomena under investigation. This comparative design examined the differential impact of feedback methods, where one group received feedback from teachers and the other received AWE feedback. The findings from this study contribute significantly to the understanding of how different feedback mechanisms influence students' emotional and motivational responses within the context of English language education.

3.3. Research Questions

This study seeks to investigate the influence of AWE tools as tools of feedback provision on students' motivation and anxiety compared to exclusively receiving teacher feedback. To accomplish these goals, the following research questions were formulated:

1. How do Automated Writing Evaluation (AWE) tools compare to teacher feedback in influencing students' motivation and anxiety levels in writing tasks?

   1a. Is there a statistically significant difference between the L2 writing motivation scores of students who received automated feedback and those who received teacher feedback?

   1b. Is there a statistically significant difference between the L2 writing anxiety scores of students who received automated feedback and those who received teacher feedback?
2. What are the perceptions of students towards the effectiveness of the two feedback modalities?

3.4. Setting

The study was carried out at the Karadeniz Technical University's English Language and Literature Department in Trabzon, Turkey, attended by students from various parts of Turkey, including students from around the world. Due to the department's primarily English-language education, students must pass a Proficiency Exam administered by the Preparatory School of Karadeniz Technical University before they can begin their studies. If successful completion of this examination is not achieved, students are required to complete a one-year English language preparatory program. As a result, for participation in ELT Department classes, students in this department must have at least a B1+ level of English proficiency. Students from this department follow a four-year curriculum and graduate with a bachelor's degree.

The course within which the study was conducted was named “Advanced Reading and Writing Skills 1” which was taught to preparatory grade students. By teaching gradual essay-writing techniques and procedures, this course seeks to improve students' written communication skills in a variety of written communication contexts and genres. Students learn about numerous writing strategies throughout the semester and develop their ability to create topic sentences, main themes, and supporting thoughts. The students are given multiple writing assignments in the form of paragraphs for which they receive feedback from their teachers. Through multiple drafts and feedback sequences, they improve and complete their essays required for the completion of the course. Once the students’ understanding of writing paragraphs has developed to a certain degree, their writing education proceeds to essays. The course also emphasizes teaching students how to develop outlines and develop arguments, which are essential skills for writing effective paragraphs.

3.5. Participants

Participants in the study were chosen from the preparatory class of the English Language and Literature Department at Karadeniz Technical University (KTU) in Trabzon, Turkey. Two primary factors served as the foundation for the reasoning behind this decision. First, the experimental and control groups in the methodological framework required splitting an already-existing classroom into two distinct units or using a classroom that was readily divided into two groups. The preparatory class was chosen because it was the only one providing a pre-existing division, especially when it came to writing training. In addition, the necessity of a writing-focused classroom that utilizes writing assignments and feedback mechanisms was met by the preparatory classroom as well. As a result, the preparatory class was chosen as the most fitting and practical for the plans of the present study (Dörnyei, 2007). Considering that students enrolled in the preparatory classes had either not achieved a passing score in the proficiency exam, which mandated a minimum English skill level of B1, or had not taken the exam at all, it could be inferred that the participants' English skill level in the study was approximately at or below the B1 level. The students' average age was predicted to

range between 19 and 22 years old, spanning the period from late adolescence to early adulthood. This age range corresponds to the period when individuals typically complete their high school education and transition to university enrollment. The study had a participation of 73 students, comprising 19 males and 54 females, which accurately represents the general population of the department, consisting largely of females. The participants were readily divided into two classrooms, which were randomly identified as two separate groups: a control group and an experimental group. The former received teacher feedback, while the latter was provided with AWE feedback. The teacher feedback group included 8 males and 29 females, while the AWE feedback group was comprised of 11 males and 25 females. The English instructor, a PhD holder who was experienced in ELT and writing instruction, also actively participated in the study as the teacher feedback provider, in addition to the students.

3.5.1. Interview Participants

Qualitative data for this study was gathered through semi-structured interviews, with participants from both the control group (N=6) and the experimental group (N=6) volunteering for the process. While the participants who volunteered to partake in the semi-structured interviews for the experimental group were 5 males and 7 females, the control group consisted of 3 males and 3 females. Interviews with the participants were conducted with all of the volunteering participants present with the condition that they had participated in the submission of the writing assignments.

3.6. Data Collection Instruments

Data collection for the study's analysis involved using multiple instruments. Specifically, pre- and post-tests as well as semi-structured interviews with volunteers were employed. In order to acquire initial insights into students' levels of writing anxiety and motivation, pre- and post-tests involving scales for writing anxiety and motivation were employed. These assessments were administered to facilitate comparisons before and after the study.

3.6.1. Demographics

To gather data and the detection of differences, similarities, and various effects among the participating students, they were asked to complete a demographic questionnaire that included factors such as age, gender, and student numbers. Students were specifically informed that their personal information would only be used for research purposes. The inquiry of student numbers had the objective of linking with their writing proficiency exam results, allowing for a comparative analysis of the two groups' writing proficiency levels prior to the start of the experimental phase.

3.6.2. Pre- and Post-tests

The research employed the Second Language Writing Anxiety Inventory (SLWAI) developed by Cheng (2004) as part of the pre-test to gauge writing anxiety levels. Given that the native language of the sample is Turkish, the current study utilized Sağlamel's (2018) translation of the inventory. This translation had undergone a thorough process of translation, cross-checking, and piloting by the researcher. Additionally, the L2 Writing Motivation Scale (L2WMS) created by Eryılmaz and Yeşilyurt (2022) was also utilized in the pre-test to assess students' motivation for writing in English (See Appendix A for the pre-test). After the experimental phase and feedback sequences, the same scales were once again administered to evaluate disparities between the control and experimental groups. The pre- and post-test questionnaires were printed on paper and provided to the students in the classroom with the addition of the pens provided complimentarily as an incentive. Paper-based questionnaires were used due to their advantages such as tangibility, availability and convenience in the classroom setting.

3.6.3. Interviews

The method of interview was chosen within this study to gain additional insight into the varying experiences of students with different feedback practices which with quantitative methods alone would be insufficient to obtain. The study particularly uses the method of semi-structured interviewing, which offers opportunities for clarifications and more detailed explanations of given responses by participants (Patton, 2002). Semi-structured interviews were conducted with volunteering student participants, which delved into their experiences, attitudes, and opinions concerning AWE tools and feedback approaches. To optimize data management for transcription and analysis, interviews were recorded using the voice recorder function available on the researcher's mobile phone. The interviews with students were focused on their individual encounters with AWE and teacher feedback. Conducted in Turkish, the native language of a majority of the participants, these discussions aimed to delve into their personal experiences, exploring their motivation and anxiety levels. Additionally, the perceived effectiveness of the AWE tool was a key topic of exploration within these interviews. Adequate informed consent was obtained, and strict measures were taken to ensure participants' anonymity and confidentiality.

3.6.4. Exam Scores

Given the readily accessible and suitably subdivided nature of the sample, there existed a potentially questionable state of normal distribution among the groups. In response to this, the preparatory exemption examinations for both groups were evaluated. The results facilitated an analysis of the differences in the preliminary academic levels of the groups. Following the assessment of eligibility for the division of groups A and B, the selected sample was advanced for further

examination. The examination scores from the preparatory exemption exam were extracted from the faculty's system for the purpose of conducting an analysis to explore any variables of the students and groups involved in the intervention. The results of this comparative analysis between the two groups are presented in Table 1 and 2.

Table 1: Exam Scores for The Groups

	Classroom	N	$\bar{X}$	S. D.
Prep exemption exam scores	A	25	41,74	16,201
	B	24	37,98	14,034

A comparison of the preliminary exam scores between the two groups was undertaken to confirm whether the performance levels were equivalent between the two groups.

Table 2: Independent Samples T-Test of Exam Scores

		Levene's Test for Equality of Variances						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Writing Scores on the prep exemption exam	Equal variances assumed	,620	,435	,867	47	,390	3,761	4,338
	Equal variances not assumed			,870	46,523	,389	3,761	4,325

The independent samples t-test was utilized to examine the writing exam scores of the groups, revealing no significant differences between the two groups ($p>0.05$). The initial similarity suggests that the groups were divided equally in terms of at the beginning of the semester.

3.7. Ethical Considerations

Before the initiation of the study, the researcher acquired approval from the KTU Social and Human Sciences Ethics Committee to ensure the ethical validation of the methods and tools utilized (refer to Appendix F). Furthermore, authorization was secured from the Department of English Language and Literature at KTU for the experimental procedures and the data collection process. The research commenced only after all necessary permissions were obtained. Participation in both the quantitative and qualitative data collection was entirely voluntary, as indicated by the textual information provided to survey participants, which clarified that their involvement was not obligatory (see Appendix A). Additionally, interview respondents were provided information regarding the process and asked for informed consent both verbally and in writing before the interview (see Appendix E).

3.8. Reliability and Validity

In quantitative research, two fundamental principles guarantee the credibility of the methods and analyses utilized: validity, which determines whether “the target construct can correctly be inferred from the particular operations of measurement” (Larry Christensen et al., 2011, p. 137), and reliability, which evaluates the “consistency or stability of scores” (Larry Christensen et al., 2011, p. 135). To ensure the reliability and validity of the quantitative instruments used in the study, a thorough evaluation process was undertaken.

Content and construct validity were established through expert reviews by specialists in English language learning and writing instruction research through critical feedback, which was instrumental in refining the instruments. The SLWAI, previously translated and validated by Sağlamel (2018), underwent further validation through his supervision to confirm its applicability in the current study. Similarly, the L2WMS was subjected to a comprehensive review to identify and address any potential issues related to content and construct validity. Preliminary reliability assessments were also conducted for both scales, demonstrating strong internal consistency. The SLWAI exhibited a reliability coefficient of 0.91, indicating high reliability. The L2WMS showed reliable internal consistency across its subscales, with Cronbach's alpha values of 0.86 for amotivation, 0.82 for intrinsic motivation, and 0.60 for extrinsic motivation. The scores confirm that the instruments used in this study are both reliable and valid, providing a sound basis for measuring the constructs of interest.

To ensure the qualitative trustworthiness of the present study, the four criteria established by Lincoln and Guba (1985)—credibility, transferability, dependability, and confirmability—were meticulously adhered to. Credibility, which parallels the internal validity criterion of quantitative research, reflects the confidence in the truth of the findings. To uphold credibility, I, as the researcher, employed a printed version of the interview transcripts, engaging in multiple readings and re-readings. Although the transcription was primarily conducted using Microsoft Word's transcription function, I meticulously reviewed the transcripts to verify their accuracy. During these multiple reviews, I conducted negative case analyses to mitigate any biases stemming from my personal expectations (Creswell & Creswell, 2017). Furthermore, to enhance both the credibility and transferability of the procedures, I provided thick descriptions of the context, sample, and findings related to the qualitative data gathering and analysis processes (Geertz, 1973). This involved establishing a detailed categorization of interview responses. To meet the criterion of dependability, a comprehensive audit trail documenting each step of the process was maintained. Additionally, methodological triangulation was achieved by incorporating multiple data sources, including both qualitative and quantitative methods, to validate the findings and achieve data saturation.

To ensure that the findings were grounded in the data and shaped by the participants rather than by my biases or expectations, I acknowledged my personal experiences with automated evaluation tools and large language models, as well as my background in writing instruction and corrective feedback. To mitigate potential biases arising from these experiences, I employed a counterbalancing approach, which included regular peer reviews. Although my expectations regarding Automated Writing Evaluation (AWE) and Artificial Intelligence (AI) tools as feedback providers were generally positive, these expectations were rigorously tested through empirical analysis within the study to ensure objective results.

The qualitative instrument used for this study was tested through piloting with one student from the same department, as well as through expert supervision. Comprehensiblity and clarity were focused on, avoiding ambiguity and focusing on one aspect at a time with each item in the interviews. Once it was assured that all items within the interview were comprehensible, clear and valid, they were proceeded with. This focus on comprehensibility clarity and what not was maintained throughout the interviews as well, by repeating questions, tone of voice, or using clarifying phrases to ensure a correct understanding of the items.

3.8. Feedback Practices and Tools

The selection of AWE tools considered the necessity for comprehensive feedback. After a thorough evaluation, PaperPal and ChatGPT were chosen due to their complementary capabilities in providing holistic feedback on both content and tone. This section intends to provide insight into the reasoning behind the utilization of PaperPal and ChatGPT as integral tools for this study. By examining the factors and considerations that informed this decision, it aims to underscore the collaborative potency resulting from their combined implementation.

3.8.1. Teacher Feedback

Being the universal standard practice of evaluating and providing feedback to student essays and writing assignments, teacher feedback was used as the feedback practice for the control group. In this study, the “feedback sandwich” method, consisting of praise-critique-praise was used to provide teacher feedback (Ferris & Hedgcock, 2005; Andrade & Evans, 2012). As the course progressed, the teacher provided feedback in the classroom as well as remotely through the Google Classroom platform on students’ paragraphs.

3.8.2. Paperpal

The decision to employ Paperpal as the AWE tool for form-based corrective feedback in this study was based on a detailed assessment of its features and applicability within the research context.

Paperpal boasts a user-friendly interface and can be easily accessed through various platforms, such as Microsoft Word integration, browser extensions, and web-based access. Notably, Paperpal's feedback goes beyond mere error identification. It offers detailed explanations that help users understand their mistakes and guides them through the correction process in real-time, fostering active learning and learner autonomy. In addition, it offers a Co-Pilot function that includes paraphrasing and shortening support, making the text more academic, and providing feedback and guidance in the form of AI chat on its paid prime subscription program. In the free-to-use version, Paperpal provides an allowance of 200 corrections per month, ensuring comprehensive feedback akin to instructor guidance. The combination of an intuitive interface, instructive feedback, and ample correction capacity made Paperpal a well-suited choice for the objectives of this study. The limitation of Paperpal was that as it announced, it “likes academic text”, which raised question marks on whether it would be able to provide feedback to beginner-level students’ paragraphs and make the appropriate corrections on their common errors, mistakes and lacks.

3.8.3. ChatGPT

Amidst its recent emergence as a globally recognized tool, the integration of ChatGPT within educational contexts has spurred substantial debates discussing its limitations, risks and benefits. The literature highlights that with educated and ethical implementation, ChatGPT has the potential to offer valuable assistance to both educators and students across various domains. Through a long period of trial-and-error analyses, and rigorous testing involving both educator-crafted content and student-authored essays, ChatGPT's proficiency in providing comprehensive content-related feedback has become evident. Despite ongoing debates regarding its educational applications, diligent testing and assessment have underscored ChatGPT's potential to deliver intricate content-oriented feedback within the realm of student-written texts. Additionally, a meticulous process of prompt engineering took place to generate a prompt that would command ChatGPT to perform in the necessary way, looking into the organization and content of student paragraphs (see appendix D for the prompt used). Throughout the evaluation process, ChatGPT was also subjected to assessments of its aptitude in supplying feedback on form-related aspects, encompassing grammar, punctuation, spelling, and word choice. Additionally, its proficiency in assigning grades or scores to student essays was examined. However, in these particular functions, ChatGPT exhibited a notably lower level of proficiency compared to alternative tools providing such services, including Grammarly, PaperPal, and Quillbot. Similar observations were made by the researcher once ChatGPT's performance was compared with other essay grading applications such as Turnitin and ProWritingAid.

3.9. Research Procedure

This study aimed to examine and differentiate the varying levels of writing motivation and anxiety among EFL students. These students received either conventional teacher feedback or

Automated Writing Evaluation (AWE) feedback on their English writing assignments. For this goal, a sample of participants was composed, consisting of preparatory-level English Language and Literature (ELL) students at Karadeniz Technical University (KTU). Through the method of convenience sampling, this sample was subsequently partitioned into two groups: the control group and the experimental group. The control group was exposed to conventional feedback methods delivered by their teachers, while the experimental group received AWE feedback. For the experimental group, a paid version of Paperpal was employed through the Microsoft Word extension to provide feedback on form aspects such as grammar, punctuation, spelling, and syntax, while the LLM ChatGPT was utilized to offer content-oriented feedback, entailing issues of coherence, relevance, and depth. Importantly, the teacher served as a moderator within the experimental group, overseeing the process and acting as a messenger between the AWE tools and the students. This teacher involvement was necessitated by considerations to ensure a balanced level of control, effective assessment, and mitigation of potential misconduct risks. The presence of the teacher within the experimental group was integral to maintaining a robust learning environment. Employing AWE tools in conjunction with teacher moderation contributed to a comprehensive feedback mechanism, enabling students to benefit from both automated and teacher insights. This dynamic approach aimed to strike a harmonious balance between the capabilities of AWE tools and the pedagogical guidance offered by instructors.

Before the experimental procedure, quantitative data was gathered through the use of the SLWAI (Cheng, 2004) and the L2WMS (Eryılmaz & Yeşilyurt, 2022) as post-tests to record students' preliminary values of motivation and anxiety regarding writing. As the first semester of the preparatory writing instruction provided students with the basic skills necessary for academic writing, students were given assignments to write paragraphs of various genres. The experimental process included a total of six assignments consisting of paragraph prompts, requiring the students to write a paragraph about the given topic via the prompts of 200 words or more. The assignments were announced and submitted via the Google Classroom platform, which is an online educational tool part of Google Apps for Education (or GAFE) that allows assignment management as well as communication between students and their teachers. Each student was allocated a period of four to seven days to submit the initial draft of their paragraphs, contingent upon the academic schedule and factors such as exams. The teacher would then provide the initial feedback within this time frame. Then, the second drafts of the essays would be expected to be submitted in 3-5 days.

3.9.1. Experimental Feedback Procedure

Across a seven-week duration, students receive a series of writing assignments along with corresponding feedback.

Table 3: Study Design

Week	Teacher Feedback Group	Automated Feedback Group
Pre-test	Pre-test (SLWAI & L2WMS)	Pre-test (SLWAI & L2WMS)
Session 1. Assignment 1	First writing assignments (Descriptive) Feedback from teacher Drafts	First writing assignments (Descriptive) Automated feedback with ChatGPT & Paperpal Drafts
Session 2. Assignment 2.	Second writing assignments (Descriptive) Feedback from teacher Drafts	Second writing assignments (Descriptive) Automated feedback with ChatGPT & Paperpal Drafts
Session 3. Assignment 3.	Third writing assignments (Process) Feedback from teacher Drafts	Third writing assignments (Process) Automated feedback with ChatGPT & Paperpal Drafts
Session 4. Assignment 4.	Fourth writing assignments (Narrative) Feedback from teacher Drafts	Fourth writing assignments (Narrative) Automated feedback with ChatGPT & Paperpal Drafts
Session 5. Assignment 5.	Fifth writing assignments (Opinion) Feedback from teacher Drafts	Fifth writing assignments (Opinion) Automated feedback with ChatGPT & Paperpal Drafts
Session 6. Assignment 6.	Sixth writing assignments (Example) Feedback from teacher Drafts	Sixth writing assignments (Example) Automated feedback with ChatGPT & Paperpal Drafts
Post-test and Interviews	Post-test (SLWAI & L2WMS) Interviews with students	Post-test (SLWAI & L2WMS) Interviews with students

An initial session, conducted one week before the assignments began, was used for the pre-test to establish baseline measurements of writing motivation and anxiety for both groups. In the following week, the first task was assigned through Google Classroom. Throghout the semester, the subsequent sessions comprised a cyclic progression of writing tasks, feedback delivery and drafts. Writing assignments of various genres (Assignments 1 to 6) were introduced, with students tasked to submit their written work, comprised of a culmination of the in-class instruction provided through the sessions. To clarify, the first assignment entailed composing a descriptive paragraph on a topic of the students' selection. The second assignment also involved crafting a descriptive paragraph, specifically focused on detailing a place the students had visited. The third assignment required the

creation of a process paragraph, aimed at explaining and elucidating a particular process. The fourth task was centered on a narrative paragraph, where students were expected to recount an event from their personal experiences. The fifth assignment involved writing an opinion paragraph, in which students were prompted to articulate their viewpoints on a specified topic. Lastly, the sixth task required the composition of an example paragraph, wherein students were to describe their perceptions of an ideal language learner. The writing feedback, enabled by AWE tools, was subsequently delivered to the experimental group, thereby fostering an iterative process of drafting and revising. In-text corrections were supplied by Paperpal, while written comments pertaining to the work were appended beneath the text by ChatGPT. Students in the control group received traditional feedback from their teachers, provided in the form of suggestions and comments under the paragraphs written by the learners. Meetings with the teacher ensured that the feedback provided by her consisted equally of both praise and criticism, while also clarifying the intervals for assignments, feedback provision, and drafts. This iterative nature of feedback aligned with the overarching aim of enhancing participants' writing quality and addressing potential areas of concern. As the study culminated in session 6, participants underwent a post-test to gauge potential shifts in their writing motivation and anxiety levels. Additionally, interviews were conducted with six volunteering participants from both groups. These interviews aimed to uncover insights into their experiences with AWE tools and their influence on writing anxiety, motivation, and performance. This sequential procedure was designed to elucidate the dynamic influence of AWE feedback on participants' perceptions, anxieties and motivation levels throughout diverse stages of the writing process. (See Appendix B for student interviews).

3.9.1.1. Prompting ChatGPT for Writing Feedback

To achieve ideal outputs from ChatGPT, a meticulous process of prompt engineering was undertaken. This involved analyzing the level, aims, and expectations of the learners in the sample. Data was gathered from meetings with the instructor, the writing assessment rubric, and tests using examples from student writings assigned before the study. A 390-word prompt was then generated for use with ChatGPT, input along with the paragraph assignments. Through multiple two-shot prompt sequences, the prompt was refined to ensure the expected output. The prompt included: (1) an introduction of the role as a feedback provider, (2) specification of the desired tone for the output, (3) an introduction to the task of evaluating and providing feedback for the paragraph entered below the prompt, (4) details of the requested output, such as praise, informed comments, corrections with specific references to the text, and evaluation of the paragraph in terms of content, vocabulary, language use, and mechanics, (5) instructions to exclude lists and bullets from the output, which ChatGPT tends to do often, and (6) a request for motivating comments and suggestions for further improvement. The resulting prompt is available in Appendix D.

3.10. Data Analysis

This research employed a mixed-method research design through a case study and gathered quantitative and qualitative data. Upon completion of the data collection phase, the collected data were systematically organized. The pre- and post-test outcomes were digitized for analysis through the SPSS software, while the recorded interviews were transcribed to facilitate thematic analysis.

3.10.1. Quantitative Data Analysis

Having gathered the results from the pre- and post-tests, the collective total of the tests was separately organized to be analysed through paired-sample t-tests through the Statistical Package for Social Sciences 27 (SPSS). To evaluate the impact of AWE feedback on students' writing motivation and anxiety, the collected data underwent an independent-sample t-test analysis. Subsequently, a paired sample t-test analysis was conducted on the combined results of both groups to identify potential differences between the pre- and post-test outcomes.

3.10.1.1. Normality Rests for Subscale analyses

For the analyses of the subscales, which displayed certain inconsistencies in normality values, the assistance of a statistician was benefited from in order to establish an accurate assessment.

To determine the appropriate method for analyzing group differences and pre- and post-test variations, Shapiro-Wilk tests were employed. For the analysis of the subscales somatic anxiety, cognitive anxiety, and avoidance behavior of the SLWAI, mixed-design ANOVAs were used, as they allowed for the simultaneous examination of both between-group and within-subject (pre- and post-test) differences. The normality of these subscales, as indicated by the Shapiro-Wilk test p-values (.428 and .092 for somatic anxiety, .082 and .248 for avoidance behavior, and .106 for cognitive anxiety pre-test), supported the use of this method. However, the cognitive anxiety post-test subscale did not exhibit normal distribution ($p = .002$). This was, however, compensated for by the skewness and kurtosis values, which were relatively low.

The data on the subscales of the L2WMS did not entirely display the same normalities as the SLWAI. The normality test p-values indicated deviations from normality for several subscales. For the intrinsic motivation subscale, the Shapiro-Wilk test p-values were .001 for the pre-test and <.001 for the post-test, indicating significant deviations from normality. Therefore, a Wilcoxon Signed-Rank test was used for this subscale. The amotivation subscale also showed significant deviations from normality, with p-values of <.001 for both pre-test and post-test, necessitating the use of a Mann-Whitney U test for analysis. Conversely, the extrinsic motivation subscale displayed normal

distribution (p-values of .200 and .370 for pre-test and post-test, respectively), permitting the use of a mixed-design ANOVA.

3.10.2. Qualitative Data Analysis

Gathered with the aim to gain additional insights into the experiences and emotions of the students, qualitative data was an integral part of this study due to its capability of offering a fine and profound understanding of the perspectives and experiences of research participants, as well as the interpretations they assign to their social reality. Braun and Clarke (2006: 79) define the aim of thematic analysis as "identifying, analysing and reporting patterns (themes) within data". As ascribed by Beuving and de Vries (2015: 157), thematic analysis is the practice of "identifying small building blocks of data and creating abstract categories from them". They also state that the method "logically flows into […] writing in naturalistic inquiry". Consequently, thematic analysis serves as an integral tool for conducting naturalistic inquiry, a research approach that seeks to investigate individuals within their everyday contexts using ordinary methodologies.

Drawing from the generic steps of thematic data analysis offered by Creswell and Creswell (2017), Figure 5 was formulated:

Figure 5: Steps of QuBalitative Analysis

The first step of the thematic analysis initiates the process by engaging in "familiarizing oneself with the data" through tasks such as transcribing when necessary, extensive reading of the data, and capturing preliminary insights (Creswell & Creswell, 2017: 87). Transitioning to step two, the process shifts to the generation of codes from "interesting features of the data" while focusing on accumulating data pertinent to each code. The third phase is characterized by the process of compiling codes into potential themes, along with the collection of all relevant data associated with each potential theme. The organization of the data continues to step four, where the researcher assesses the relation of the themes and coded extracts by "generating a thematic map of the analysis" (Creswell & Creswell, 2012: 87). Step five sharpens the focus of the analysis on the process of "defining and naming themes" (Creswell & Creswell, 2017: 87). Concluding the journey, step six synthesizes findings through a finalizing assessment "of selected extracts" with regard to the aims and background of the research.

The researcher carefully reviewed and read the transcriptions multiple times to "gain an overall sense of the information and reflect on its overall meaning" (Creswell & Creswell, 2017: 21). The first step involved listening to the recordings and transcribing them word by word into text. This required listening to the recordings multiple times to correctly understand each word, grasp the content and become familiar with various aspects of the data. Once the transcripts were thoroughly reviewed, the researcher proceeded to assign codes and categorize similar or repetitive data patterns into segments. Any particular instances of keywords or statements generated were highlighted, considering their relevance to the themes. In conducting the qualitative analysis, as the researcher, I acknowledged my personal experiences with automated evaluation tools and LLMs, and my background in writing instruction and corrective feedback. To minimize potential etes arising from these experiences, I employed a counterbalancing approach throughout the study, including regular peer reviews. Although my expectations of AWE and AI tools as feedback providers were generally positive, these expectations were systematically tested through empirical analysis within the study to ensure objective results.

This chapter demonstrated the design, setting, participants, data collection instruments and procedures, and the analysis of the gathered data of the present research. Employing a mixed-method design, this research benefited from both quantitative data in the form of exam scores and questionnaires and qualitative data through observations throughout the process and semi-structured interviews. Following the pretest with the entire sample of the study which consisted of 73 preparatory grade students, the experimental phase provided teacher feedback to Group A (N=37) and automated feedback through ChatGPT and Paperpal to Group B (N=36). Noteworthy functions and abilities along with weaknesses of the tools were observed and noted throughout the procedure. Following the process that comprised multiple drafts and instances of feedback through a total of six tasks, the pretest was applied. The last instance of data collection took place in the form of semi-structured interviews with volunteering participants from both groups (N=12), aiming for a fair and accurate representation. Analysis of qualitative data was conducted through thematic analyses, while quantitative analysis in the form of descriptive and inferential tests was performed using the IBM SPSS software and the researcher aimed to discover answers to the research questions initially formulated for the study.

CHAPTER FOUR

4. FINDINGS AND DISCUSSION

4.1. Introduction

In this chapter, the study's findings are discussed along with an explanation of their interpretation. A description of the results of each data collection method is followed by a presentation of the quantitative findings, including exam results and a questionnaire measuring learners' writing motivation and anxiety.

4.2. Quantitative Findings

This section will present the quantitative data finding of the present study. Results of the analyses will be presented along with tables.

4.2.2. Writing Anxiety

The analysis of data acquired via the SLWAI indicated minimal significant differences in anxiety levels between the two groups. However, both groups exhibited a reduction in anxiety levels. Further examination of the subscales suggested encouraging results, demonstrating a notable improvement in cognitive anxiety in both the groups. Despite these factors, the employment of automated feedback specifically demonstrated no significant variance in anxiety levels among the groups upon being assessed through the Second Language Writing Anxiety Inventory (SLWAI).

4.2.2.1. Second Language Writing Anxiety Inventory (SLWAI)

Designed to measure the writing anxiety levels of second language learners of English with subgroups of avoidance behaviour, somatic anxiety, and cognitive anxiety, the SLWAI was chosen and used to measure and categorise the participating students' writing anxiety levels.

4.2.2.2. Reliability of the SLWAI

An analysis was performed to confirm the reliability of the SLWAI scale, demonstrating promising results, with a Cronbach's alpha of ,916 for the pre-test and ,924 for the post-test. The

results of the normality tests for the variables indicated that if the p-value exceeded 0,05, the distribution could be considered normal (Ghasemi & Zahediasl, 2012). Notably, only the cognitive post-test yielded a p-value lower than 0,05. However, given that both skewness and kurtosis values are minimal, a de facto normal distribution can also be assumed for this variable. Consequently, these variables were deemed appropriate for the analysis using an analysis of variance (ANOVA).

Table 4: The Basic Psychometric Properties of the SLWAI Scale

Scale		$\bar{X}$	S. D.	Skewness	Kurtosis	Normality test p value	Cronbach's Alpha
SLWAI	Pre-test	2,96	0,72	-0,74	-0,35	,960	,916
	Post-test	2,73	0,74	-0,30	-0,23	,236	,924
Somatic	Pre-test	2,96	0,92	0,12	-0,65	,428	,878
	Post-test	2,77	0,88	0,43	-0,44	,092	,888
Avoidance	Pre-test	2,58	0,74	-0,03	-0,90	,082	,746
	Post-test	2,47	0,75	0,32	-0,17	,248	,780
Cognitive	Pre-test	3,35	0,92	-0,33	-0,53	,106	,873
	Post-test	2,93	0,94	-0,53	-0,70	,002	,882

Note. The Shapiro-Wilk test was used to test the normal distribution.

Table 5: T-test Resultsof the Pre-test of SLWAI

		N	$\bar{X}$	S. D.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
1	I feel anxiety while writing in English.	73	3,03	1,105	9,6	23,3	28,8	31,5	6,8
2	I feel anxious while writing in English under time constraint.	73	4,03	,942	1,4	4,1	21,9	35,6	37,0
3	While writing in English, I feel worried if I know they will be evaluated.	73	3,78	1,083	4,1	11,0	13,7	45,2	26,0
4	I often write down my thoughts in Turkish and translate them into English.	73	2,18	1,206	35,6	34,2	12,3	12,3	5,5
5	I usually do not write in English unless it is compulsory.	73	2,88	1,258	17,8	23,3	20,5	30,1	8,2
6	My thoughts become jumbled when I start writing in English in an exam.	73	2,75	1,362	21,9	27,4	17,8	19,2	13,7
7	I worry that my English writing is worse than that of my classmates.	73	3,14	1,475	19,2	19,2	15,1	21,9	24,7
8	If my English composition is to be evaluated, I would worry about getting a very poor grade.	73	3,96	1,136	2,7	12,3	12,3	31,5	41,1
9	I do my best to avoid situations in which I have to write in English (e.g., social media).	73	1,81	,967	47,9	30,1	17,8	1,4	2,7

Table 5: (Continued)

		N	X̄	S. D.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
10	My thoughts become jumbled when I write in English under time constraint.	73	3,47	1,179	4,1	21,9	19,2	32,9	21,9
11	I prefer to use English to write.	73	3,51	1,132	6,8	9,6	30,1	32,9	20,5
12	I am afraid that the others would deride my English writing if they read it.	73	2,67	1,365	23,3	30,1	16,4	16,4	13,7
13	I freeze up when unexpectedly asked to write English compositions.	73	3,23	1,253	13,7	8,2	38,4	20,5	19,2
14	I would do my best to avoid writing in English.	73	2,21	1,054	28,8	35,6	26,0	5,5	4,1
15	I don’t worry at all about what other people would think of my English writings.	73	3,04	1,296	15,1	19,2	28,8	20,5	16,4
16	I am afraid of my English writing being chosen as a sample for discussion in class.	73	3,47	1,313	11,0	13,7	19,2	30,1	26,0
17	I usually seek chances to write in English outside of class.	73	2,59	1,153	20,5	30,1	21,9	24,7	2,7
18	When writing in English, I usually feel tense.	73	2,78	1,181	13,7	34,2	19,2	26,0	6,8
19	I am afraid that my English paragraphs would be rated as very poor.	73	3,56	1,225	8,2	13,7	15,1	39,7	23,3
20	I can't think of anything when I start doing writing assignments.	73	2,56	1,130	15,1	41,1	24,7	11	8,2
21	I don’t feel nervous about showing my writings to my peers.	73	2,79	1,213	16,4	28,8	20,5	27,4	6,8
22	I start trembling when I am writing in English.	73	1,74	1,106	58,9	21,9	9,6	5,5	4,1

The anxiety levels of both groups appeared to be moderately high for the pre-test, as revealed by the statement on the first item “I feel anxiety while writing in English” which displayed a mean score of 3,03 for the entire sample, with 31,5% agreeing and 6,8% strongly agreeing with the statement. The level of anxiety is heightened by the introduction of time constraints as revealed by the responses for the statement “I feel anxious while writing in English under time constraint.” (4,03), with anxiety related to time constraints highlighted when compared to the same statement without time constraints. A similar theme appears with the statement “My thoughts become jumbled when I write in English under time constraint.” (3,47).

Anxiety about being assessed as poor or receiving a poor grade seemed to be present among both groups. For being rated as poor, a relatively high mean score was detected for the statement “I am afraid that my English compositions would be rated as very poor” (3,56) and for receiving a poor

grade, with the statement “If my English composition is to be evaluated, I would worry about getting a very poor grade.” (3,96). Although not as high, a moderately high level of anxiety about doing worse than classmates was also present in the post-tests of all participants (3,14). This hesitance is also present in situations where one’s writing is viewed by or compared within the classroom, with moderately high means for statements “I am afraid of my English writing being chosen as a sample for discussion in class” (3,47) and “I worry that my English writing is worse than that of my classmates.” (3,14).

The willingness to write in English does not appear to be uniformly low across contexts, with a mean of 2,56 for the statement 'I usually do not write in English unless it is compulsory', while a moderately higher mean of 3,51 for the statement 'I prefer to use English to write' indicates the presence of a willingness to write. An analysis of the items suggests that while writing is preferred as a social or leisure activity rather than as a compulsory educational activity, compulsory writing seems to evoke an anxious, avoidant reaction, as reflected in statements such as “I freeze up when I am unexpectedly asked to write English compositions”, which has a moderately high mean (3,23). The statement “I do my best to avoid situations in which I have to write in English (e.g., social media).” also displays a lower, much lower mean of 1,81, supporting this notion. Physical responses to writing and reactions also revealed moderately high anxiety levels. A higher mean score for the statement “I freeze up when unexpectedly asked to write English compositions.” (3,23) and (What else?) reveals that students experience high anxiety in in-class writing circumstances, although a low mean score for the statement “I start trembling when I am writing in English”. have also been observed (1,74). Additionally, the relatively high mean score for “My thoughts become jumbled when I write in English under time constraint” (3,47) further underscores the stress associated with writing under pressure in the form of time constraints.

Interaction with peers also appears to have been demonstrated as a cause of anxiety, with the item “I worry that my English writing is worse than that of my classmates”. having a relatively high mean score of 3,14 as well as the item “I don’t feel nervous about showing my writings to my peers”. demonstrated a relatively low mean (2,79).

Table 6: T-test Results of the Post-test of SLWAI

		N	**$\bar{X}$**	**S. D.**	**Strongly disagree**	**Disagree**	**Neutral**	**Agree**	**Strongly agree**
1	I feel anxious while writing in English.	66	2,59	1,189	17,8	30,1	19,2	17,8	5,5
2	I feel anxious while writing in English under time constraint.	66	3,76	1,053	4,1	8,2	13,7	43,8	20,5

Table 6: (Continued)

		N	$\bar{X}$	S. D.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
3	While writing in English, I feel worried if I know they will be evaluated	66	3,29	1,237	9,6	17,8	12,3	38,4	12,3
4	I often write down my thoughts in Turkish and translate them into English.	66	2,23	1,134	26	35,6	16,4	6,8	5,5
5	I usually do not write in English unless it is compulsory.	66	2,56	1,217	20,5	28,8	15,1	21,9	4,1
6	My thoughts become jumbled when I start writing in English in an exam.	66	2,82	1,189	11	31,5	19,2	20,5	8,2
7	I worry that my English writing is worse than that of my classmates.	66	2,71	1,333	21,9	21,9	15,1	23,3	8,2
8	If my English composition is to be evaluated, I would worry about getting a very poor grade.	66	3,58	1,177	6,8	13,7	6,8	46,6	16,4
9	I do my best to avoid situations in which I have to write in English (e.g., social media).	66	1,95	1,182	42,5	26	11	5,5	5,5
10	My thoughts become jumbled when I write in English under time constraint.	66	3,45	1,192	4,1	19,2	19,2	27,4	20,5
11	I prefer to use English to write.	66	3,2	1,07	5,5	15,1	38,4	19,2	12,3
12	I am afraid that the others would deride my English writing if they read it.	66	2,32	1,23	28,8	27,4	16,4	12,3	5,5
13	I freeze up when unexpectedly asked to write English compositions.	66	2,7	1,277	15,1	34,2	15,1	15,1	11
14	I would do my best to avoid writing in English.	66	2,12	0,886	23,3	38,4	24,7	2,7	1,4
15	I don’t worry at all about what other people would think of my English writings.	66	2,82	1,3	17,8	21,9	19,2	21,9	9,6

Table 6: (Continued)

		N	$\bar{X}$	S. D.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
16	I am afraid of my English writing being chosen as a sample for discussion in class.	66	3,03	1,312	15,1	17,8	19,2	26	12,3
17	I usually seek chances to write in English outside of class.	66	2,61	1,201	17,8	27,4	26	11	8,2
18	When writing in English, I usually feel tense.	66	2,42	1,151	20,5	34,2	16,4	15,1	4,1
19	I am afraid that my English compositions would be rated as very poor.	66	3,11	1,349	15,1	17,8	13,7	30,1	13,7
20	I can't think of anything when I start doing writing assignments.	66	2,71	1,16	11	35,6	20,5	15,1	8,2
21	I don’t feel nervous about showing my writings to my peers.	66	2,64	1,223	19,2	23,3	27,4	12,3	8,2
22	I start trembling when I am writing in English.	66	1,56	0,93	57,5	23,3	2,7	5,5	1,4

On the post-test, students exhibited lower levels of anxiety about writing in English, with a mean score of 2,59 for the statement, “I feel anxiety while writing in English”. However, higher anxiety rates under time constraints remained an issue, with a mean score of 3,45, and only a 0,02 decrease in the mean score compared to the pre-test. This may be attributed to the writing examinations employed to assess and grade students, which were chosen to prevent them from using AI tools to complete assignments. While assignments were still used, students were not graded based on the content of their writing but instead based on the submission of assignments. This heightened response rate in comparison to the pre-test for time constraints and writing activities may suggest that students' anxiety levels were elevated by the pressure of deadlines and exams. In addition to the previously mentioned anxiety related to time constraints, the sample also exhibited heightened anxiety regarding evaluations (M=3,29). This may suggest that receiving scores for their writing assignments and being assessed on their writing skills might have increased students’ anxiety levels.

There has been a notable decline in students' fears of their writings being evaluated, with the mean score dropping from 3,78 to 3,29. In contrast, while still low, interest in writing in situations where one might be expected to write in English, such as social media, has increased, with the mean score rising from 1,81 to 1,95. Additionally, a positive shift has been observed in the perception of poor grades. This is evidenced by the change in the mean score for the statement, “If my English

composition is to be evaluated, I would worry about getting a poor grade," which decreased from 3,96 to 3,58. This suggests that students may be becoming more familiar with the writing class and improving their writing skills, concurrently reducing their anxiety levels.

The preference to write in English also appears to have decreased from a mean of 3,51 to 3,20. However, the level of anxiety related to being asked to write in class unexpectedly does not display a similar level of anxiety compared to the pre-test. The mean for the statement "I freeze up when unexpectedly asked to write English compositions" significantly decreased from 3,23 to 2,70. The level of anxiety related to compulsory writing for in-class purposes has decreased from 2,88 to 2,56. This suggests a reduction in reluctance to write in English for in-class assignments.

Assignment-related anxiety has also increased, as evidenced by the mean score of the statement "I can't think of anything when I start doing writing assignments" which has risen to 2,71 from the initial 2,56. Although there has been a reduction in the fear of being rated poorly from 3,56 to 3,11 on the mean scores, the statement "My thoughts become jumbled when I start writing in English in an exam" has exhibited a slight elevation from 2,75 to 2,82, suggesting an increase in cognitive anxiety related to examinations and the organisation of thoughts. Tendency to write in Turkish and then translate into English also seems to have increased slightly, from a mean of 2,18 to 2,23. While this practice is not conducive to the development of metacognitive skills in English writing, it appears that students were challenged by the expectations set for the tasks and chose to translate rather than write.

Generally, anxiety in the classroom setting also seems to have increased for the entire sample, with the mean score for the statement 'I can't think of anything when I start writing assignments' increasing from 2,56 to 2,71. Although the fear of being marked poorly seems to have decreased from a mean score of 3,56 to 3,11, the statement 'My thoughts become confused when I start to write in English in an exam' seems to have undergone a slight increase from 2,75 to 2,82. Anxiety related to being compared to others and others' thoughts on one's writing has appeared to increase, with the mean of "I don't worry at all about what people think of my English writing" falling from 3,04 to 2,82. This may have been influenced by the feedback sequences and the amount of correction and praise in the feedback as well as in classroom practice. However, the anxiety that one's writing is inferior to that of classmates has decreased from 3,14 to 2,71. Similarly, the fear of being judged and/or ridiculed by others for one's mistakes has decreased from 2,67 to 2,32. In a similar vein, the anxiety of being chosen as a sample for class discussion has also decreased from 3,47 to 3,03. It is possible that this type of anxiety was reduced as a result of becoming familiarised with classmates and through the gradual improvement of one's skills.

4.2.2.3. Findings of the Mixed-Design ANOVA

In order to measure differences between groups as well as differences between pre and post-test results, the mixed design ANOVA has been used due to its ability to analyze elements of both time, as in pre and post-test results, and groups such as experimental and control groups. Given that reliability analysis ensured that the data was eligible to be tested via an ANOVA, this method of analysis was chosen to be the most fitting for the current study.

Table 7: Somatic Anxiety Mixed-Design ANOVA Results

	df	F	p	Partial ETA Squared
Between Groups				
Group	1	0,01	,916	,00
Within Groups				
Time	1	3,50	,066	,05
Groups x Time	1	0,40	,531	,01

An analysis of variance with a mixed design, examining the somatic subscale of the anxiety inventory and comparing the teacher-provided feedback group to the automated feedback group across pre-test and post-test measurements, indicates no statistically significant differences between the groups or in the pre-test and post-test results for somatic anxiety. This conclusion is supported by a P-value above 0,05, suggesting that the somatic anxiety levels of the participants have not altered to a statistically significant extent.

Table 8: Avoidance Behaviour Mixed-Design ANOVA Results

	df	F	p	Partial ETA Squared
Between Groups				
Group	1	0,04	,841	,00
Within Groups				
Time	1	1,96	,167	,03
Groups x Time	1	3,80	,056	,06

The examination of the avoidance subscale, utilizing the same analytical method to compare differences between groups and pre-post test outcomes, demonstrates a similarly high P value, exceeding 0,05. Although there appears to be a reduction in avoidance behavior, declining from a mean of 2,58 to 2,47, the data collected do not reveal any significant or discernible differences between the two groups under study.

Table 9: Cognitive Anxiety Mixed-Design ANOVA Results

	df	F	p	Partial ETA Squared
Between Groups				
Group	1	0,00	,955	,00
Within Groups				
Time	1	31,36	<,001	,33
Groups x Time	1	1,13	,291	,02

The data suggests that for cognitive anxiety, pre-test scores are generally higher than post-test scores, with the pre-test scores exhibiting a mean of 3,35 and a standard deviation of 0,92, compared to the post-test scores, which display a mean of 2,93 and a standard deviation of 0,94. This favourable trend in cognitive anxiety, observed across both participant groups, indicates a progressive decrease in anxiety levels as students become acclimatized to the course demands and enhance their writing abilities through constructive feedback.

4.2.3. Writing Motivation

In this section, the results of the quantitative data obtained using the L2WMS scale, which includes pre-tests conducted before and post-tests conducted after the experimental phase for both the control and experimental groups are presented. Additionally, this section includes an assessment of the findings from the scale, focusing on the differences observed between the control and experimental groups based on outcomes of the pre and post-test. Moreover, the reliability analysis of the L2WMS is also detailed herein.

4.2.3.1. L2 Writing Motivation Scale (L2WMS)

An assessment of students' writing motivation levels and whether their motivation was intrinsic or extrinsic was carried out using data gathered via the L2WMS.

4.2.3.2. Reliability of the L2WMS

Table 10: The Basic Psychometric Properties of the L2WMS Scale

Scale & Subcale		$\bar{X}$	S. D.	Skewness	Kurtosis	P Value	Cronbach's Alpha
L2WMS	Pre-test	2,86	0,39	-0,42	0,34	,211	,618
	Post-test	2,80	0,45	0,35	0,05	,139	,586
Intrinsic Motivation	Pre-test	3,94	0,69	-0,95	1,88	,001	,735
	Post-test	3,75	0,84	-1,18	2,24	<,001	,837
Extrinsic Motivation	Pre-test	2,99	0,72	-0,12	-0,15	,200	,587
	Post-test	2,86	0,84	0,29	-0,14	,370	,723
Amotivation	Pre-test	1,33	0,50	1,54	1,57	<,001	,652
	Post-test	1,52	0,84	2,20	5,23	<,001	,874

Before proceeding with an analysis of the subscales and differences between the control and experimental groups, the L2WMS was put through a reliability analysis. The majority of Cronbach's Alpha values exceeded the threshold of 0,70, with only two falling below this level. Values that approximate this threshold are generally considered adequate (Ghasemi & Zahediasl, 2012). Should the p-value associated with the variables surpass the value of ,05 in a normality test, this indicates that they conform to a normal distribution, thereby rendering them appropriate for analysis using ANOVA. The results of the L2WMS subscale in general, and the subscale of Intrinsic Motivation, on the contrary, do not adhere to a normal distribution. Consequently, to analyze the variable, it is imperative to utilize non-parametric tests such as the Mann-Whitney U Test and the Wilcoxon Signed-Rank Test, rather than ANOVA (Pallant, 2020). Extrinsic Motivation, which is normally distributed, is suitable for evaluation using ANOVA. Conversely, Amotivation, which does not follow a normal distribution, necessitates the application of non-parametric tests, specifically the Mann-Whitney U Test and the Wilcoxon Signed-Rank Test, in place of ANOVA.

Table 11: T-test Scores for the Pre-test of L2WMS

		N	V	S. D.	Str on	Dis ag	Ne utr	Ag ree	Str on
1	Since I reveal my ego/myself (while writing in English), my desire to write in English increases.	73	3,88	,957	1,4	8,2	19,2	43,8	27,4
2	Since writing in English improves my abilities, my desire to write in English increases.	73	4,16	,782	0	5,5	6,8	53,4	34,2
3	Because I get satisfied with and enjoy learning new things, my desire to write in English increases.	73	4,12	,849	0	5,5	13,7	43,8	37
4	Since I want to show myself that I am a successful person, my desire to write in English increases.	73	3,66	1,216	8,2	9,6	17,8	37,0	27,4
5	Since I enjoy writing in English, my desire to write in English increases.	73	3,77	1,007	2,7	6,8	27,4	37,0	26,0
6	When I fail the English course, my desire to write in English increases.	73	2,56	1,291	28,8	20,5	23,3	20,5	6,8
7	When I attend courses, my desire to write in English increases.	73	3,85	1,076	4,1	6,8	20,5	37,0	31,5
8	Because my family wants me to be an English literate, my desire to write in English increases.	73	2,23	1,349	41,1	24,7	12,3	15,1	5,5
9	When I'm away from technology, my desire to write in English increases.	73	2,37	1,087	27,4	26,0	30,1	15,1	1,4
10	Because I think it will help me find a good job, my desire to write in English increases.	73	4,10	1,082	5,5	4,1	8,2	39,7	42,5
11	To be honest, I don’t know. I really think I'm wasting my time while writing in English.	73	1,23	,514	79,5	19,2	0	1,4	0
12	I have no idea. I do not understand what I'm doing when I'm writing in English.	73	1,33	,746	78,1	15,1	4,1	1,4	1,4
13	Writing in English is far from me.	73	1,67	,973	57,5	26,0	11,0	2,7	2,7
14	I do not find writing in English meaningful.	73	1,21	,499	82,2	16,4	0	1,4	0

The preliminary findings derived from the analysis of the L2WMS scale suggest that the participants demonstrate a notably high level of motivation, as evidenced by a mean pre-test score of 3,88. The initial data highlight a robust motivation among participants to utilize writing as a means of expression, a fact accentuated by the foremost item on the scale. Moreover, the highest mean score on the scale, recorded at 4,16, points to a motivation predominantly driven by the motivation to improve personal writing skills. This elevated score indicates that participants acknowledge the significant advantages offered by writing courses in advancing their personal development. Furthermore, a mean score of 4,12 on the item stating "Because I get satisfied with and enjoy learning new things, my desire to write in English increases" supports the claim that the initial motivation to write among participants is considerably high. In contrast, the item "Since I want to show myself that I am a successful person, my desire to write in English increases" gained a rather lower average score of 3,66. This could suggest that the motivation to write is primarily driven by a genuine interest in writing and enhancing writing proficiency, rather than by the need to affirm personal success or demonstrate competence.

Despite presenting a slight discrepancy, the motivation to write, driven by an intrinsic interest in the act of writing itself, recorded an average score of 3,77. Although this score is somewhat lower than those of the preceding items, it remains relatively elevated. The marginally lower score for this specific item may imply that while learners are motivated to improve their skills through writing, their motivation may not be as pronounced when it pertains to writing for personal enjoyment. Failure or a low grade in writing courses does not appear to be a motivator to write for the participants, which displayed a relatively lower mean score of 2,56 which highlights a disconnection between academic failures and motivational responses. Nevertheless, attendance to courses displays a higher mean score of 3,85, which reflects that the instruction provided in courses increases students' motivation to write along with their skills. Family expectations seem to exhibit a minimal influence on writing motivation, as indicated by a relatively low mean of 2,23. In contrast, the potential for improved job prospects for individuals possessing advanced writing skills plays a substantial role, evidenced by a higher mean of 4,10, underscoring its significance in motivating writing. Moreover, the low mean of 2,37 regarding writing in the absence of technology implies that the participants, likely due to their status as digital natives, demonstrate a pronounced preference for using technological tools like computers or mobile devices for writing, which facilitates digital input.

The levels of amotivation observed in the study were notably low, with a mean score of 1,23, reflecting a substantial majority of participants who rejected the notion that learning English writing is a waste of time. Additionally, the data indicated that the inability of participants to comprehend their activities during English writing sessions was similarly minimal, with a mean score of 1,33. This suggests a significant level of awareness among the participants regarding both the processes and the benefits of writing instruction. Furthermore, the statement "Writing in English is far from me" received a relatively low mean score of 1,67, underscoring a limited disinterest in English

writing within the sample. This finding is corroborated by the response to the statement “I do not find writing in English meaningful,” which recorded the lowest mean score of 1,21, indicating a strong recognition among participants of the value of writing in English.

Table 12: T-test Scores for the Post-test of L2WMS

		N	$\bar{X}$	S. D.	Strongly Disagree	Disagree	Neutral	Agree	Strongly agree
1	Since I reveal my while writing in English, my desire to write in English increases.	65	3,71	1,182	6,8	8,2	12,3	38,4	23,3
2	Since writing in English improves my abilities, my desire to write in English increases.	65	4,03	,984	2,7	5,5	8,2	42,5	30,1
3	Because I get satisfied with and enjoy learning new things, my desire to write in English increases.	65	3,85	,956	2,7	6,8	11,0	49,3	19,2
4	Since I want to show myself that I am a successful person, my desire to write in English increases.	65	3,45	1,199	8,2	12,3	15,1	38,4	15,1
5	Since I enjoy writing in English, my desire to write in English increases.	65	3,72	1,038	5,5	4,1	17,8	43,8	17,8
6	When I fail the English course, my desire to write in English increases.	65	2,57	1,212	17,8	30,1	21,9	11,0	8,2
7	When I attend courses, my desire to write in English increases.	65	3,51	1,017	4,1	9,6	24,7	38,4	12,3
8	Because my family wants me to be an English literate, my desire to write in English increases.	65	2,22	1,375	42,5	12,3	12,3	16,4	5,5
9	When I'm away from technology, my desire to write in English increases.	65	2,35	1,280	31,5	19,2	19,2	13,7	5,5
10	Because I think it will help me find a good job, my desire to write in English increases.	65	3,68	1,200	8,2	6,8	12,3	39,7	21,9
11	To be honest, I don’t know. I really think I'm wasting my time while writing in English.	65	1,46	,920	65,8	12,3	5,5	4,1	1,4
12	I have no idea. I do not understand what I'm doing when I'm writing in English.	65	1,55	,952	58,9	19,2	4,1	5,5	1,4
13	Writing in English is far from me.	65	1,62	1,085	58,9	17,8	4,1	4,1	4,1
14	I do not find writing in English meaningful.	65	1,45	,985	68,5	11,0	2,7	4,1	2,7

The post-test results reveal a slight reduction in the motivation levels among participants, particularly in their motivation to engage in writing activities. The initial item on the motivation scale, specifically addressing the motivation to write, displayed an average score of 3,71, indicating a subtle decrease in motivation. Although this decrement suggests a potential waning in students' motivation, it is important to note that motivation levels remained relatively high. Similarly, a slight decline was noted in the response to the statement, “Since writing in English improves my abilities, my desire to write in English increases,” where the mean score decreased from 4,16 to 4,03. Additionally, the motivation to write, influenced by the satisfaction derived from acquiring new knowledge, also demonstrated a minor reduction, moving from a mean of 4,12 to 3,85. This decline in motivation could be partially attributed to participants' displeasure with the writing prompts and topics provided during the study. A notable decrease was also observed in those who had indicated agreement or strong agreement with the item measuring self-affirmation of success as a motivator

(*M*=3,45) on the pre-test. While 64,4% of the participants had either agreed or strongly agreed with this item, this percentage was decreased to 53,5% on the post-test.

A willingness to write for the enjoyment of it, while also experiencing a slight decrease (*M*=3,72), remained relatively stable for the post-test. With enjoyment remaining as a consistent motivator, failure in English courses remains virtually unchanged with only a minor leap from 2,56 to 2,57, suggesting that most of the participants still disagree that exam failure would increase their motivation to write. However, a notable decline is evident for the statement When I attend courses, my desire to write increases" (*M*=3,51), indicating a diminishing return on motivation from attendance at writing courses. Expectations of family members remain stable at a mean of 2,22, suggesting that family expectations continue to exert a limited influence on participants' willingness to write in English. Similarly, the majority of participants indicated a low level of motivation to write in the absence of technology, with a mean score of 2,35, which had only decreased by 0,05 compared to the pre-test. This suggests that distancing oneself from technology does not motivate writing.

In examining the impact of job opportunities on perceptions of writing skills, a significant decrease in the mean score to 3,68 was observed, suggesting a diminished valuation of advanced writing skills for career advancement. Although still markedly low, there is a slight increase in the average perception of writing instruction as wasteful, with scores rising from 1,33 to 1,55. The consistent and low mean score of 1,62 for the perceived irrelevance of writing instruction could indicate a stable recognition among participants of the importance of enhancing their English writing skills. Furthermore, the statement "I do not find writing in English meaningful" registered a mean score of 1,45, indicating a minor increase but still reflecting that the majority of the sample finds writing in English to be meaningful.

4.2.3.3. L2WMS Subscales

Following the subscales divided by Eryılmaz and Yeşilyurt (2022), which facilitated the assessment of (1) intrinsic motivation, encompassed by items 1 through 5; (2) extrinsic motivation, covered by items 6 to 10; and (3) amotivation, represented by items 11 to 14, the L2WMS was investigated on the different categories of motivational orientations exhibited by the participants. Due the the aforementioned P-value results of the reliability analysis, Mann-Whitney U tests were used for the Intrinsic Motivation and Amotivation subscales.

Table 13: L2WMS Intrinsic Mann-Whitney U Test

	Control Group		Experimental Group			
	X̄	**Std. Deviation**	**X̄**	**Std. Deviation**	**p value**	**U statistics**
Pre-test	3,93	0,73	3,91	0,63	,718	633,50
Post-test	3,74	0,77	3,76	0,91	,501	477,00

A Wilcoxon Signed-Rank test was employed to compare the Intrinsic Motivation scores obtained in the pre and post-tests, separately for both the control and experimental groups. It was determined that the difference in the intrinsic motivation levels of the teacher feedback group, as measured in the pre-test and post-test, was not statistically significant, with a p-value that is slightly above the 0,05 threshold on the teacher feedback group. Similarly, the automated feedback group exhibited no significant difference in the Intrinsic Motivation scores between the pre-test and post-test evaluations. The findings underscore a statistically insignificant disparity between the groups, with the teacher feedback group demonstrating slightly more favourable outcomes.

Table 14: L2WMS Extrinsic Mixed-Design ANOVA

	df	F	p	Partial ETA Squared
Between Groups				
Group	1	0,00	,962	,00
Within Groups				
Time	1	2,34	,131	,04
Group x Time	1	1,56	,217	,02

Note. The Shapiro-Wilk test was used to test the normal distribution.

In the analysis conducted using a mixed-design ANOVA, which simultaneously compared classes and test times, the L2WMS Extrinsic Motivation subscale revealed no significant differences among the groups. This outcome is attributed to the fact that all groups exhibited p-values exceeding the threshold of 0,05. Consequently, it may be deduced that the extrinsic motivations exhibited by both the teacher and automated feedback groups have remained comparably stable, without significant variations to distinguish either group.

Table 15: L2WMS Amotivation Wilcoxon Signed-Rank Test

	Pre-test		Post-test			
	$\bar{X}$	S. D.	$\bar{X}$	S. D.	P vale	Z statistics
A Sınıfı	1,36	0,49	1,36	0,63	,908	-0,12
B Sınıfı	1,35	0,51	1,67	0,99	,014	-2,45

For the automated feedback group, the Wilcoxon Signed-Rank test revealed a moderate statistical significance difference between the amotivation scores from the pre-test and the post-test, evidenced by a p-value of 0,014, which is below the threshold of 0,05. Conversely, for the teacher feedback group, the comparison between the pre and post-test amotivation scores did not demonstrate a statistically significant difference. The findings of this analysis of amotivation levels reveal that the amotivation levels of the experimental group have experienced a somewhat larger increase when compared to that of the control group.

4.3. Qualitative Findings

Qualitative data were utilised in the current study to establish a complex and holistic understanding of the affectional positions as well as the various perspectives of the participants in a natural setting (Creswell & Poth, 2016). The interview transcripts, produced as word-for-word transcriptions of the recorded conversations of the semi-structured interviews with participants, were analysed using the content analysis method, which was determined to be the most appropriate approach for the present study, enabling a complete and thorough review of the qualitative data (Miles et al., 1994). Students' interactions with teachers and automated feedback were explained and contextualised through an analysis of interview data from a representative subset of the sample, allowing the research questions to be addressed in part. The findings of the thematic analysis conducted on the interview transcripts are presented and discussed in this section.

4.3.1. Anxiety

*"At this point, I don't want to make the teacher say or even think "It's the end of the year and you're still making these (mistakes)?"" (*Respondent A3M)

The qualitative findings on anxiety gathered through individual interviews with voluntary respondents were partly in conflict with the quantitative data. A detailed investigation of the respondents' statements and attitudes revealed distinct similarities and contrasts between the control and experimental groups.

Table 16: Participants' responses regarding writing anxiety

Participant	Excerpt
A1F	*When I get feedback, I'm like "hmm, so I have a lot of mistakes", so it might be breaking my self-confidence a bit.*
A3M	*At this point I don't want to make the teacher say or even think "It's the end of the year and you're still making these (mistakes)?*
A4M	*There was more tension, part of the time was spent thinking "what can I do" - I can't do much good during that stress.*
A5F	*The teacher (feedback), you know - you feel a bit embarrassed when you get it wrong*
A6M	*At first, there is inevitably the anxiety and the thought of "can I do it?", "will I write it?, "what will I write?".*
B1M	*I was a little taken aback by the higher level of grammatical errors than I expected.*
B2F	*I feel stressed because I have shortcomings, making mistakes and writing incorrectly makes you nervous.*
B3M	*I was quite surprised when at first, I wondered if I hadn't written it well.*
B4F	*Because I know that I make mistakes, I may have difficulties in writing because I lacked that self-confidence at the beginning.*
B5F	*At first, when I got feedback, I was a bit nervous upon opening (the file) initially. Then, when I opened it, sometimes, most of the things were crossed out in red. For example, I saw the mistakes I made.*
B6F	*I was really upset when I saw things that I was sure were true being crossed out. I was like, "Am I that bad?"*

While the control group regularly mentioned the impact of feedback on their self-efficacy and confidence, participants in the experimental group more frequently reported that their anxiety was reduced by the recognition and correction of their errors. For example, A3M refers to the presence of pressure that comes with a fear of making mistakes in the writing process, as well as the fear of disappointing the instructor by failing to improve his writing skills, which may be illustrative of a negative outcome of teacher feedback leading to anxiety through the pressure of self-criticism and responsibility to the instructor. Furthermore, he mentioned experiencing stress as a result of lacking insight into how to precisely correct his mistakes, a problem for which he sought help from ChatGPT. On the other hand, B4F pertains to an initial anxiety and stress that existed at the beginning of the writing process but later faded away with the help of the feedback, and her mistakes were rectified. B5F made similar remarks, beginning with the words ”at first” while discussing her anxiety, which was alleviated by the process of receiving feedback and addressing her mistakes. She stated that she felt anxious before examining the feedback, but she stated that once she viewed the feedback file, it was ”not a problem”. Similarly, B2F stated that she was experiencing high levels of anxiety at the beginning of the semester, but these levels were reduced by seeing and correcting her mistakes.

Anxiety was mentioned in relation to the assignments and processes of thinking through writing and submitting. Respondent A3M described a scenario in which he felt increasing stress upon encountering a new assignment as he wondered what to write for the task and how he could write a text long enough to be submittable. He also mentioned that he is most challenged when faced with situations where he must improvise or work under pressure. Students’ responses frequently hinted at the notion that negative, late, and unexplained feedback, along with writing prompts that were outside their area of interest, were reasons for anxious attitudes toward writing. The majority of the respondents (n=7) reported lower levels of anxiety related to their writing being examined and assessed in front of the classroom. Four respondents in the experimental group said it would not bother them, while three respondents in the control group said they would not be anxious when it came to being assessed in front of the class. Anxiety about making mistakes, not being able to identify their mistakes, and not being able to fix and eventually reduce the number of mistakes in the writing process were prevalent themes overall. In the experimental group, interviewees regularly reported their anxiety in the past tense, indicating that, while they had initially been anxious, this had decreased significantly.

These incidents demonstrate the crucial role of comprehensive feedback, straightforward explanations for corrections, and consideration of students' development, as well as recurring errors, in reducing learners' writing anxiety. It is also important to note that although similar incidents occurred in the control group, they were not as frequent or vivid as those in the experimental group.

Table 17: Participants' responses about anxiety dissolving

Participant	Excerpt
A3M	*I'm happy when I see that I'm making fewer mistakes. Without this, my stress probably wouldn't have decreased.*
B1M	*When I saw that the number of mistakes was high, I can say that I became even more ambitious.*
B2F	*After I understood things. You know, this (anxiety) has decreased.*
B3M	*An anxiety settled in there. After that (feedback), this anxiety turned into motivation. From motivation, it continued in a positive way.*
B4F	*After a while, you become a little more eager because you realise that you are getting better and better through feedback.*
B5F	*At first, it stresses me out before opening it, "I wonder what they wrote about me", but after opening it, there is no problem afterwards*
B6F	*Yes, I was very hesitant to write before I got feedback, but after the feedback, I was thinking "I should write and improve even more*

Feedback and self-efficacy were among the factors that influenced the level, type, and frequency of writing anxiety. The pressure of receiving feedback and a loss of self-confidence were often mentioned by participants in the control group (A), particularly when they felt they had made mistakes. For example, A1F claimed that feedback made her feel less confident and A5F expressed regret for her mistakes. A3M highlighted that he felt nervous when he was given tasks one after the other, and that he was worried about the challenges ahead. When queried regarding the evolution of his perspectives on writing anxiety, he articulated that there had been a notable decrease in his writing anxiety. However, he did not attribute this reduction to the positive impacts of the feedback received from the teacher. Rather, he emphasized his personal endeavors as the primary factor in mitigating his anxiety. Furthermore, he suggested that the onus of alleviating such anxiety ultimately resides with the students themselves. Similarly, A6M described some initial anxiety and uncertainty before starting the task. The participants in the automated feedback group (B) expressed similar concerns. B2F noted that her fear of making mistakes and her perceived shortcomings caused her stress. B4F showed how the feedback had a positive impact on her motivation and expressed her relief upon seeing improvements after the initial anxiety. The majority of participants indicated that their anxiety did not significantly impact their writing performance. Many participants, especially those in the automated feedback group, reported that their anxiety was alleviated through feedback, as it allowed them to identify, correct, and gradually reduce the number of errors in their writing. Noteworthy examples within the automated feedback group include respondent B2F, who generally exhibited an anxious demeanour, and B6F, who often expressed a dismissive attitude towards writing instruction. While both groups experienced reduced anxiety, the experimental group reported a higher frequency of anxiety reduction and more students acknowledging these changes. Conversely, several respondents in the experimental group indicated that their anxiety had diminished, vanished, or been replaced by an eagerness to write more, improve their skills, or review the feedback file.

Anxiety and motivation were frequently discussed in relation to one another, with the notion suggesting that, either through time, the student's own efforts, or with the assistance of the feedback

process, anxiety was eventually lessened or replaced by motivation. Throughout the interviews with the respondents, positive responses regarding this dissolution of anxiety and the emergence of motivation were more prevalent with the automated feedback group, although similar responses were provided by those in the control group as well.

4.3.2. Motivation

"At that point, all of a sudden, anxiety came over me. Then, this anxiety turned into motivation. And from that motivation, it continued positively."

Participant B3M

Owing to the topics asked during the interviews and the participants' frequent mentions of how particular factors either favourably or adversely affected their desire to write, motivation frequently appeared throughout the interviews. Anxiety and motivation were often referred to adjacently, stating that, either through time, the student's own effort, or with the help of the feedback process, anxiety was either reduced to some extent or eventually replaced by motivation.

Table 18: Responses on motivation

Respondent	Excerpt
A1F	*Positive feedback is already the biggest factor (for motivation)*
A2F	*I try to look at this situation positively because it makes me feel good since I have the chance to see my mistakes.*
A3M	*I am also happy when I see that the mistakes are decreasing, my self-confidence increases*
A4M	*I think motivation will increase more if there is a friendly environment.*
A5F	*If it is like "this is good, but this can be improved" instead of just "you did this wrong, you did that wrong", of course my motivation will improve.*
A6M	*the fact that someone values the things we write, that is, someone reads it and tells you your mistakes, motivates the student more.*
B1M	*I used to scroll directly to the bottom of the page and read there (the comments) rather than the crossed-out parts, as they were the parts that were more valuable to me because of the direct spot-on determinations made there. Therefore, I considered it to be a very good method.*
B2F	*(Feedback) affected me, I mean it affected me positively. I saw my mistake, I corrected it, I said "this is the right one.*
B3M	*The comments and feedback stimulated a desire for improvement. This desire for improvement was also supported by motivation*
B4F	*Small notes like "happy writing", though they are minor, I think they soften the language and increase your willingness.*
B5F	*(Feedback) makes me more ambitious and enthusiastic, I like to correct my mistakes and rewrite that paragraph again.*
B6F	*You wrote each one with personal consideration, that's why it was better, you explained it and you wrote messages like "keep it up" which made me more willing.*

In the comparative analysis of feedback mechanisms within the study groups, it was observed that the automated feedback group expressed a notable increase in motivation, predominantly

attributed to the positive reinforcement embedded in the feedback process. Participants such as A1F highlighted that praise concerning the strengths of her writing significantly boosted her motivation. Similarly, B4F commended the inclusion of encouraging phrases like “happy writing” or “keep writing” in the feedback generated by ChatGPT, which was deliberately prompted to add a layer of praise. Furthermore, B6F appreciated the personalized nature of the feedback, which acknowledged individual efforts, thereby enhancing the perceived value of the feedback received. This sentiment was echoed by A5F, who emphasized the importance of feedback that not only identifies errors but also celebrates successful elements in writing, underscoring the benefits of constructive criticism. The efficacy of personalized feedback was further supported by other participants in the automated feedback group, such as B1M and B3M, who appreciated the tailored comments accompanying their written paragraphs. They noted that this approach not only motivated them but also cultivated a sense of having an engaged audience for their work. This collective feedback underscores the critical role of positive reinforcement and personalized engagement in enhancing motivation and overall participant satisfaction with the feedback process.

4.3.3. Intrinsic and Extrinsic Motivation

Although the quantitative data on extrinsic and intrinsic motivation proved to be somewhat ambiguous, the differences between respondents' statements regarding their types of motivation helped explain their experiences and engagement with the feedback they had received. Both groups presented a blend of both forms of motivation within this classification, but the source of motivation displayed certain differences. The teacher feedback group frequently acknowledged the teacher's authority and interpersonal interactions between students and feedback providers, as well as the validating role of the teacher, whereas the experimental group seemed to appreciate their personal growth, improvement in writing skills, and reduction in errors. For instance, participants A1F, A2F, and A4M underscored the motivating factor of positive feedback, that is, positive comments on the strong aspects of their writing. As revealed earlier, the anxiety of failing to satisfy the teacher’s expectations and improve the self is also reflected in A3M’s responses regarding motivation, where he states that he wishes to reduce the number of mistakes in his writing. As per B6F's report, her social activity, the drama group, had prevented her from devoting time to writing; however, she reported that the feedback had motivated her to write and improve, and that she had developed a curiosity about how she could develop herself after receiving the feedback. Two participants from the AWE feedback group, B2F and B5F, also shared their positive experiences with the feedback process. B2F mentioned that reviewing her mistakes in the feedback helped her gain a better understanding of how her writing skills had developed. Meanwhile, B5F expressed a motivation to write more frequently, as she perceived her writing to be improving. Additionally, B5F stated that she would be delighted to receive writing assignments on a daily basis. The feedback provided by ChatGPT and Paperpal seems to have had a positive impact on motivation, especially for participants such as B2F, B4F, and B5F, who stated that they were not intrinsically motivated to write or were

not motivated at all to begin with. This seems to be a sign that the feedback provided by ChatGPT and Paperpal is at a level that can compete with, or at least coexist with, the feedback provided by the teacher.

4.3.4. Students' ChatGPT use for writing

The majority of the participants acknowledged ChatGPT, while a relative number stated that they did not use it. It was observed that the use or knowledge of ChatGPT was viewed negatively by several participants who either hesitated or strongly denied any use or knowledge of ChatGPT. A5F, who admits to using ChatGPT to complete assignments, states that she received "some" help from ChatGPT to complete one of her assignments, adding the excuse that she was ill. Although he was assured of anonymity and informed that his statements would be used solely for research purposes, it is uncertain why B3M confidently denied any ChatGPT use for assignments, whether they used paraphrasing tools such as QuillBot to avoid detection, or whether he was concerned about certain repercussions related to his AI use. In order to prevent discomfort, he was not confronted about his detected assignment. Two noteworthy exceptions, particularly because they were within the control group, were A3M and A4M, who stated that they used ChatGPT for their writing assignments, but not to generate content, but interestingly to receive additional feedback.

Table 19: Participants who used ChatGPT

Participant	Excerpt
A3M	*In writing assignments, I can copy and paste my mistakes in the form of "tell me where I made a mistake", which is also very convenient.*
A4M	*It is true that I use (ChatGPT) in a "can you explain what's wrong here" way, which I think is more normal. If the teacher is inaccessible or takes a long time, I think this is normal.*
A5F	*I once wrote a paragraph with a little help from ChatGPT, but I was sick."*
B1M	*I use ChatGPT in the same way as Grammarly, either to correct minor mistakes or in terms of correcting grammar, as I said, formal writing. But from time to time I don't know exactly how to put the subject in my mind into text. I can get help from ChatGPT in this regard.*
B3M	*At that time, I got a lot of help from this tool ChatGPT and Grammarly, but then I realised that these tools are also flawed. I realised that artificial intelligence is also flawed.*

While the prohibitions and preventative measures, as well as the unfavourable view of the use of language models in education, may be justified because of their unethical use for content generation and completion of tasks with little educational growth and fairness, the specific way in which participants A3M and A4M utilised ChatGPT could be inculcated into learners to establish an environment of healthy AI habits that promotes learner autonomy and further benefits both educators and learners by assigning AI tools a more constructive role rather than one where they are shunned and are better avoided.

Table 20: Participants' comments about ChatGPT

Participant	Excerpt
A1F	*The lecturer's feedback make me feel like they contain a more personal, real human thought.*
A3M	*(Teacher feedback) also makes me feel a little more confident. After all, the teacher is a person I meet, a person I know. No matter how good the website is, even if it speaks beautifully there, in the end, the teacher feedback feels like it is the best.*
A4M	*ChatGPTs cannot know us. Maybe if we tell them about ourselves, it can give more accurate feedback.*
A5F	*I would be less embarrassed. I would be less embarrassed when I make a mistake, you know, when I can't do something. Other than that, with the teacher, you feel a little embarrassed when you make a mistake.*
B1M	*Whether it is an artificial intelligence (providing feedback), it shows that there is a value given to that text.*
B3M	*To understand if the other person is a human being, they need to relate to you. A human being would usually say "I also have such and such a thing" and give examples from themselves, but (ChatGPT) cuts the praise part right there.*

Several participants in the experimental group expressed a preference for teacher feedback and authority of the teacher figure in the classroom over feedback from GenAI. However, they had already received a sequence of feedback provided by ChatGPT and Paperpal, which, although was moderated by the researcher, had successfully disguised itself as teacher feedback. The experimental group also described AI feedback as potentially repetitive, dishonest, superficial, and lacking insight into development and recurring mistakes. Nevertheless, they praised the feedback they received as being beneficial, original, motivating, and concise. The decision to not initially disclose to the participants in the experimental group that they would receive AI feedback was based on a prediction that seems to have been confirmed to prevent bias and the influence of preconceived notions from clouding their judgment. It was also aimed at ensuring that students would not look for flaws in the AI-generated feedback, which they would be less likely to do with feedback from a teacher.

4.3.5. Writing Prompts/Subjects

Almost all the interview participants referred to the significant effect of the writing subjects they were assigned, stating that they affected their motivation to write and anxiety, which weakened their ability to write.

Table 21: Comments about writing topics as a motivator

Respondent	Excerpt
A1F	*The teacher gives a topic in the lesson, it matches exactly with what I have in mind. Then I feel more motivated."*
A2F	*Anxiety happens when I can't find a topic or when I feel stuck . When I do not have much knowledge on that subject, or when I cannot write.*
A3M	*I mean, for example, a difficult assignment comes up, or I don't know, a subject that I don't have a good grasp of. Then I feel this kind of anxiety.*
A4M	*When it's about something I've never experienced before, you know, I can get stressed.*

Table 21: (Continued)

Respondent	Excerpt
A5F	*I mean, I don't feel any (motivation) until I find the topic. But after I find the topic, I write, I produce".*
A6M	*If the subject does not appeal to me, I can get a little more nervous.*
B2F	*I think it would make more sense if we wrote about something we want to write about, rather than a specific topic.*
B3M	*The title and topic of the assignment affect me a lot. If the topic is, as I said, very extreme and something I haven't thought about, I get quite stunned.*
B4F	*I can be more willing if there is a subject that interests me, because it is a subject that I have mastered and I like.*
B5F	*It doesn't cause me any stress, especially when there are interesting topics.*
B6F	*The lecturer gives a topic in class, it matches exactly what I have in mind. Then I am more motivated.*

A clear agreement across both groups indicates that the nature of writing assignments has a significant effect on their motivation to write and anxiety of writing, whether they align with their personal interest or are perceived as relevant to their knowledge or that of the general population. Respondents A1F and A6M highlighted the positive effects of topics that resonate with their personal interests or ideas, thus increasing their motivation. On the other hand, A2F, A3M, and A4M expressed concerns about the challenges posed by topics that were either unfamiliar or uninteresting, leading to an increased sense of anxiety and a tendency to finish the task with minimal engagement. A5F's remarks highlight the initial demotivation faced when struggling to find a suitable topic, although this changed to a more positive outlook when an engaging topic was encountered. On the other hand, B2F and B6f underlined the role of free-choice topics, implying that self-selected topics could lead to greater motivation. B3M's and B4F's experiences reinforce this sentiment, with personal interest in a topic serving as a stimulus for improved performance and motivation. B5F's statement expands on the binary of stress levels, depending on whether the topic falls within her areas of interest, highlighting how topic relevance can reduce or increase writing anxiety. The analysis suggests that the automated feedback group may have a stronger desire for autonomy, which may indicate a tendency towards personalised writing practice. In any case, it is clear that these observations highlight the need for more relevant writing topics.

4.3.6. Preferred method and type of feedback

The student participants engaged in the study were required to complete writing assignments during class sessions, as part of their coursework, and in examination settings. Feedback was provided to them in various forms: directly during class on their handwritten assignments, and digitally through the Google Classroom platform. Having been exposed to diverse methods and approaches in both writing tasks and receiving feedback, which included both instructor and automated feedback, the participants displayed distinct preferences regarding the types of feedback they favored.

4.3.7. Face to face feedback

It has been frequently noted by students from various groups that there is a strong preference for face-to-face feedback from instructors. This method not only facilitates a deeper explanation of corrections but also enhances the detail and directive nature of the feedback through more expansive commentary. Certain participants have highlighted face-to-face feedback as a preferable solution to address issues related to the timeliness of feedback delivery. B4F stated that face-to-face feedback tends to be more enduring and comprehensible compared to online feedback, which she finds less memorable. She further suggests that transforming written comments into voice memos could bridge the gap between online and face-to-face feedback, potentially rendering the feedback more effective and personal. Additionally, A1F points out that receiving feedback in a classroom setting not only allows for more detailed feedback but also facilitates learning from the mistakes of peers, thus enriching the educational experience.

Table 22: Respondents' comments about face-to-face feedback

Respondent	Excerpt
A1F	*When we meet face to face with the teacher in the classroom, her feedback seems reliable to me. ChatGPT is an automatic tool after all. Its feedback felt very ordinary and unoriginal to me, but the teacher's feedback felt more like an individual, real human thought.*
B2F	*Face to face feedback, for example in the classroom with the teacher, I think it is more effective. Because, for example, I don't know if we can write to you, but I don't understand when you give feedback, for example, I don't know if I can convey this to you.*
B4F	*If you are asking this question in terms of learning, the feedback given face to face is more memorable and more understandable. When it's only on text, can lose its memorability and comprehensibility.*
A6M	*I think it is much better to read and evaluate the assignments in class. Because the mistake made by one person is often also made by someone else, and when someone else sees it in class, it can be more effective.*

Participants in the teacher feedback group expressed a clear preference for face-to-face feedback, emphasising the value they saw in the personalised, human elements of such interactions. A1F highlighted the reliability and authenticity of face-to-face feedback from a teacher, which she perceived as more individualised and reflective of genuine human thought than the more generic use of language in language models. This preference emphasises the importance of establishing a personal connection and ensuring that feedback is customised to an individual's specific needs and work. While participants in the automated feedback group expressed a similar liking for face-to-face feedback, their comments also revealed preconceptions about the limitations of automated systems in providing feedback. B2F raises practical concerns about the process of receiving feedback from automated systems, questioning whether it allows for real-time clarifications and dialogue about misunderstandings. This relates to the perceived lack of interactivity and responsiveness in the AWE tools. B4F supports this view by pointing out that face-to-face feedback is generally more memorable and understandable than written feedback, which might suggest that the presence of a human

interlocutor can enhance the comprehension and retention of feedback, likely due to the dynamic nature of face-to-face communication.

4.3.8. Positive Feedback

All respondents mentioned that they would greatly appreciate positive feedback, that is, feedback that praises the positive or accomplished aspects of their writing. This is most often referred to as a method of motivating learners and counteracting the demotivation associated with corrections. This is also reiterated by A2F, who states that it would also help learners know what they have done correctly and should continue to do in their writing. Positive comments and suggestions, as well as praise for parts where students did well, were also included in the automated feedback prompt, where the computer stated how the student had successfully explained a part of their writing. This was done in the experimental group to match the praise-criticism-praise feature of the feedback sandwich approach, which was also acknowledged by the teacher who provided feedback for the control group.

Table 23: Respondents views about praise in feedback

Respondent	Excerpt
A1F	*Of course, it should be in a way to emphasise the flaws, but if the good sides are also given, people can actually be motivated with effective feedback.*
A2F	*In fact, good comments could be made on top of considering the mistakes. An example would be "It's good that you touched on this subject, but you could do that other sentence like this"*
A5F	*If instead of just "you did this and that wrong", it said "this is good, but this can be corrected", of course my motivation would improve.*
B4F	*Although our weaknesses are guided while receiving feedback for our development, the good aspects or the words, the language we use an the feeling we add should also be explained clearly, or some details such as "I like that part" could be added.*

The impact and importance of delivering positive feedback and complimentary comments along with criticism of learners' mistakes and pointing out areas for improvement were apparent throughout the interviews, as participants, particularly those in the teacher feedback group, revealed that such feedback could impact their writing development as well as their motivation to continue writing and to participate in writing classes. Statements made by those in the teacher feedback group were often phrased in words such as "could" or "would" in the form of suggestions or wishes. In contrast, while participants in the AWE feedback group did not refer to a preference for positive feedback and praise, participants such as B1M and B3M pointed out that receiving feedback made them feel valued and recognised through a sense of audience.

4.3.9. Feedback on content

The decision to choose ChatGPT along with Paperpal was due to its ability to analyse and provide suggestions, particularly on the organisation, argument, flow, and overall content of writing.

Feedback regarding content, making suggestions on how to improve the flow, organisation of both ideas and sentences, arguments, explanations, and narrations, were highly praised, and students were more than happy to have received such feedback. While those in the experimental group praised this aspect of the feedback they received, participants in the control group raised their dissatisfaction in terms of the lack of feedback on the content aspects of their paragraphs. The comments provided by ChatGPT, which were copied and pasted at the bottom of student paragraphs, appeared to have created a beneficial experience for the students to see where they could improve their writing in terms of content. A lack of feedback regarding the content of the assignments was also frequently reported by the participants in the teacher feedback group.

4.3.10. Written feedback in the form of comments

Comments on areas for improvement, areas of ambiguity, methods for correcting mistakes, and areas of strength in a student's writing constitute a significant component of valued feedback. Respondents in the experimental group particularly appreciated the inclusion of positive and motivating comments at the beginning of feedback paragraphs, which consisted of one or two sentences highlighting the strengths of the students' writing. This approach was designed to introduce a layer of praise before transitioning to constructive criticism. The feedback would then integrate the praise and criticism within a single sentence and conclude with a motivational phrase such as “keep writing” or “well done”. It is evident that students were encouraged by the presence of such affirming words and phrases, which aimed to boost their motivation without discouraging them from recognizing their errors. Respondent BF6 remarked that the “personal consideration” embedded in each comment enhanced the overall experience, and the positive remarks further motivated her. Likely understanding how much time it would take to individually read and write such detailed feedback for the assignments of every single student for a teacher, she asked “isn’t it very time consuming to send feedback to everyone?”.

4.3.11. Detailed Explanations

Detailed explanations of what aspects of one’s writing should be changed, what are the exact language features to work on, how and why certain words, phrases and even sentence are crossed out, elements in one’s writing that not be present or elements that need to be changed are a critical part of a complete corrective feedback for writing. These findings on the significance of detailed explanations of written corrective feedback align with those of Bruno and Santos (2010).

Table 24: Respondents' comments about detailed explanations

Respondent	Excerpt
A1F	*I couldn't respond to most of the feedback because I didn't know what to fix, so my performance score is very low. I just can't, because there's nothing to fix. I don't know what to correct.*
A2F	*feedback could be more detailed, but I don't know how the features of the feedback should be, so..*
A3M	*The teacher tells us our mistake, and I solve it with ChatGPT, for example, perhaps explanations like "Here is your mistake", or "you should write like this" could work.*
A4M	*When they say "you have grammatical mistakes", we sometimes have to read the whole text from top to bottom. This can be a problem especially for those who write long texts.*
A5F	*In terms of grammar, for example, I get feedback saying "pay attention to your grammar", but if I knew what was wrong with it, I would have done it correctly when I first wrote it.*
A6M	*When the teacher gave the feedback, I was surprised when she said "correct grammatical mistakes", you know, "was there a mistake?". Then, when I looked at it again and read it 2-3 times, I said "Aha, yes".*
B1M	*The fact that there were separate suggestions and comments for everyone's text makes you feel that what you have put forward has been recognised.*
B3M	*They were actually quite good. They could make almost all the points. You could probably say that I paid the most attention to them. I was looking at them first before the corrections.*
B5F	*For example, if there is a subject that I am missing, or if there is a subject that is too much or if it deviates too much from the subject, I find it absolutely sufficient because it explains it exactly.*
B6F	*It was very nice, because it felt very special so from each one was really neatly answered and done personally*

A comparative analysis of the statements made by participants in both groups revealed a statistically significant discrepancy in the perception of feedback. Those in the teacher feedback group frequently expressed frustration due to the lack of clarity in the explanations provided, which they felt hindered in their ability to comprehend and rectify their errors. By contrast, participants in the automated feedback group did not raise any concerns regarding the feedback they received.

The automated feedback group consistently commended the feedback for its thoroughness and effectiveness in helping students identify and correct their mistakes. This approach was perceived as "spot-on" and sufficiently supportive to avoid discouraging students (B1M). Participant B5F, for instance, stated that she was content with the number of comments and explanations provided in the feedback but also expressed a preference for more comments on additional topics to be included in her writing assignments. For instance, she stated on multiple occasions that she would benefit from suggestions on additional topics for discussion in her paragraphs. While the majority of students in the experimental group receiving automated feedback expressed satisfaction with the level of detail present in the comments provided to their assignments, a few participants in the experimental group have also added that they would appreciate more detailed explanations regarding certain corrections or suggestions (B1M; B6F)

Table 25: Timeliness of feedback

Respondent	Excerpt
A1F	*We still do writing in the lessons, but since there is no immediate checking of the assignments at that moment, our responsibility is a bit too much. Maybe if they could be more immediate during the lesson, the feedback would be more noticeable.*
A3M	*For example, the feedback of a very old assignment would be given just now, sometimes I even forget what I wrote, so when it is that much, the writings do not make much sense for me.*
A5F	*Well, it has an effect like this, now if it (the feedback) comes during school time, we already have done it, but when it comes on holiday, I don't feel like doing it and I didn't do it.*
B6F	*If the teacher gives feedback to each of us at that moment, this is probably a bit too much to expect, but it would be easier for us. Because when we don't do it at that moment, when we go home and do it, it's gone, my mind leaves that place. I don't want to do it again.*

Several participants frequently mentioned the issue of timeliness in the teacher feedback group and one participant in the AWE feedback group, although respondents were not specifically questioned about the speed of feedback. While the majority of the students in the automated feedback group did not mention the timeliness of the feedback as an issue, the students in the teacher feedback group frequently mentioned that feedback could be timelier or that an alternative practice would be preferable, such as receiving feedback in the classroom. Similarly, while students in the automated feedback group also mentioned that it would be preferable to receive feedback in the classroom, this preference may be due to the fact that in-class teacher feedback is provided in real time and continuously as the student writes. One potential explanation for this discrepancy may be that in in-class teacher feedback, the provision of feedback is instantaneous, whereas online feedback or feedback on assignments takes longer to reach the learner. However, the example of B6F suggests that in situ teacher feedback within the classroom may always be necessary. This issue has been raised on numerous occasions by students as a justification for the shortcomings of their feedback. A4M proposed that ChatGPT can be employed to address the ambiguity of feedback that the recipient may not comprehend without further clarification. ChatGPT can be employed by students for independent studies and to receive feedback from AI, thereby enabling them to enhance their writing skills autonomously.

In the teacher feedback group, some students were more fixated on the issue of timely feedback, with A5F iteratively raising the issue of not receiving feedback quickly enough. She added that this contributed to her missing corrections, being unable to write and submit assignments, and concluding with a low score for participation. From the experimental group, B6F, who had previously stated that she found it difficult to devote time to assignments, stated, similarly to A3M, that she became disconnected from and unfamiliar with the writing she had done in the classroom once she returned home and received feedback on it, adding that she would prefer in-class feedback to address this issue, as she did not wish to return to the same piece of writing to work on it afterwards. With this in consideration, regardless how timely writing feedback is provided either by teacher or technological

tools, students with lower levels of engagement and motivation to self-study would still prefer instantaneous feedback in class. An alternative to B4F stated that online feedback could take the form of video or voice recordings, helping with comprehension with the help of stress and intonation in the instructor's voice.

4.4. Conclusion

This section offers a concise overview of the conclusions and insights derived from this study. It seeks to elucidate the impact of both automated and teacher feedback on the writing anxiety and motivation levels of EFL learners. Furthermore, this part of the present thesis illuminates how these findings can enhance and optimize writing instruction and facilitate the integration of automated feedback tools and AI technologies within language education contexts.

The current investigation provides a broad examination, incorporating both qualitative and quantitative methods, of the multifaceted ways in which feedback techniques can be related to the levels of writing anxiety and motivation. However, it is important to recognize that the findings of this study should be considered as exploratory. They highlight the need for further research into the affective responses of language learners to different feedback methods. There is also an urgent need to investigate the impact of automated feedback via genAI on learners' long-term progression.

The present study was designed with these principal objectives in mind: firstly, to examine the impacts of two feedback modalities—namely, teacher feedback and automated feedback— provided through prompted responses of ChatGPT, and the writing assistance tool Paperpal, particularly in terms of their effects on learners' motivation and anxiety levels within the context of EFL; secondly, to investigate the connection between students emotional reactions to writing tasks and feedback, specifically in terms of the factors of motivation and anxiety; thirdly, to evaluate the correlation between said emotional dimensions and to what extent they were changed through two different kinds of feedback. The overarching goal of these objectives consisted of enhancing the comprehension of the connections between psychological and instructional factors, and to offer insights that could help create more efficacious methodologies in future studies and practices.

The SLWAI was utilized in order to gather quantitative insights into the anxiety levels of the students within the sample of preparatory grade students of the department of English language and Literature program at Karadeniz Technical University. Preliminary findings of the inventory revealed moderately high levels of anxiety with 38,3% of the participants reporting a moderate or strong agreement with feeling anxious while writing in English, with this anxiety particularly increased through the introduction of time constraints. Being assessed as poor or being compared with peers appeared to be issues related to the moderately high anxiety levels, similar to the findings of previous studies on anxiety related to being assessed (Aydın, 2008). The research identifying the substantial

impact of time constraints on anxiety levels among EFL learners echoes the conclusions of earlier studies, such as those conducted by Ekmekçi (2018) as well as Türnük and Aydın (2020). The present study similarly highlighted time constraints as notable contributors to anxiety among Turkish EFL students. The observed levels of moderate preliminary anxiety overlap with those reported by other scholars within the Turkish or Middle Eastern EFL setting, such as Kayaoğlu and Sağlamel (2013), Aydın (2012) and Sevinç (2018).

Following the conclusion of the experimental phase, wherein the experimental group was exposed to feedback from ChatGPT and Paperpal, moderated by the researcher, and the control group received conventional feedback, also through Google Classroom, from the teacher on their writing assignments submitted through Google Classroom, the post-test of the Second Language Writing Anxiety Inventory (SLWAI) was conducted. The results of this post-test revealed a notable decrease in the anxiety levels across the entire sample, with 47,9% of the participants disagreeing or strongly disagreeing with feeling anxious about writing in English. Despite the general reduction, anxiety related to time constraints remained unchanged, with 64,3% of the entire population agreeing or strongly agreeing with feeling stressed when writing under time constraints. Although anxiety associated with evaluations or receiving poor grades exhibited a significant decrease, the method of composing text in the native language and subsequently translating it into English, rather than writing directly in English, has become increasingly prevalent, as evidenced by comparisons with pre-test results. The practice of composing text in one's native language and subsequently translating it into a target language is advised against, as direct composition in the target language has been suggested to be the most efficacious method for EFL learners (Cohen & Brooks-Carson, 2001). For those with lower levels of proficiency at English writing, however, using a certain level of translation from the native language could prove to be a beneficial method to use before transitioning not directly writing in the target language (Huh, 2001).

In terms of anxiety related to writing assignments, the post-test revealed a slight increase in the entirety of the sample, further highlighting the impacts of time constraints and traditional writing instruction methods on learners' anxiety levels. Fear of being mocked by others for one's writing displayed a decrease, likely affected by familiarization with peers and the gradual improvement of the learners' own skills. Similarly, it is probable that through frequent exposure to writing exercises within the educational environment, the occurrence of anxiety relate to students being unexpectedly asked to write during class sessions diminished markedly, from an average of 3,23 to 2,70. A discernible decrease in general anxiety levels among the participants was observed; however, anxiety specifically related to examinations remained constant. Additionally, there was an increased tendency to initially draft text in the native language and then translate into English, as opposed to composing directly in the target language. These findings contribute to a nuanced understanding of the anxiety associated with writing and receiving feedback.

Upon the analysis of the quantitative data to ascertain the presence of any noticeable differences between the two groups, the findings indicated an absence of significant disparities in the values of somatic anxiety and avoidance behavior across the groups. Importantly, a significant reduction in cognitive anxiety was observed between the pre-test and post-test periods; however, this reduction was consistent across both groups. This temporal difference, coupled with the parallel outcomes observed in both groups, suggests that both teacher feedback and automated feedback contributed positively to the reduction of cognitive anxiety levels within the sample.

The L2WMS was utilized in order to gather quantitative data regarding the participants motivation levels to write in English and with the aim of discerning differences indeed intrinsic and extrinsic motivation levels along with the lack thereof of motivation among and between the control and experimental groups. The moderately high levels of anxiety present within the groups were accompanied by a relatively high level of writing motivation, as evidenced by 71,2% of the participants agreeing or strongly agreeing that they were motivated by the ability to express themselves in English. The predominant motivator for writing, as evidenced by the highest average score on this scale, was the widely held belief that writing enhances overall English proficiency. This sentiment was corroborated by more than 87% of participants who indicated their agreement or strong agreement with the notion that their English skills were significantly improved through the act of writing. The participants motivation to write exhibited a strong inclination toward the category of intrinsic motivation, supported by the agreement with items testing intrinsic motivation, accompanied by lower mean scores on those that measure extrinsic motivation. One notable deviation from the norm of extrinsic motivation is the significant influence of the potential for securing successful job opportunities, which has emerged as a compelling motivator for writing in English. This is evidenced by 82,2% of the participants affirming or strongly affirming its impact. Preliminary results from the pre-test reveal a robust and positive classification of motivation among the participants, driven by personal growth, enjoyment, and career benefits, as opposed to external factors or dependencies. This motivation profile underscores the necessity for interventions aimed at enhancing motivation to write in English or improving writing skills to focus on intrinsic motivation levels and professional aspirations. Such interventions should provide learners with materials and tasks that cater to their specific motivational needs.

Following the intervention, motivation levels of the entire sample entire sample dropped which displays an overall decrease in the willingness and enjoyment of the participants. Particularly, motivation to write as an expressive means underwent a slight decrease from 3,88 to 3,71. The mean score that reflected writing motivation fed by the perception that it helps develop one's English skills also reduced to 4,03 from the initial 4,16 on the pre-test. The observed decline in intrinsic motivation among participants may be attributed to various factors including the nature of the writing instruction, the materials utilized during the sessions, or the themes and genres of the writing assignments assigned to the students. This trend is partially elucidated through qualitative data analysis, which

indicates a consistent pattern: the intrinsic motivation levels of the participants decreased, likely as a result of the mandatory nature of the writing tasks and instruction. This hypothesis is further corroborated by the relatively stable mean score of 3,72 on the item that measures the participants' motivation to write merely for the enjoyment of writing in English. Conversely, the impact of course attendance on writing motivation seems to have diminished, suggesting a potential decline in the motivational influence of the instructional environment over time. Despite these trends, the motivational effects of external factors, such as career opportunities, remain relatively unchanged with slightly lower mean scores, although the agreement with the motivating role of career prospects has only experienced a minor decline.

In conclusion, the comparative analysis of pre and post-tests clearly indicates a discernible decline in intrinsic motivation and the enjoyment of writing. However, it is important to note that the intrinsic motivation levels of the participants experienced merely a slight diminution, primarily associated with instructional influences and self-improvement efforts. These observations underscore the importance of implementing engaging writing activities and assignments. Furthermore, they call for a thorough reevaluation of educational approaches and methods in writing instruction. Such strategic enhancements are crucial for sustaining and potentially increasing the motivation for writing among foreign language learners.

In order to discern differences between the groups and between pre-test and post-test measurements, an analysis of the subscales was conducted along with a comparative analysis of the two groups' responses. The results of these analyses indicated that the responses between the two groups regarding intrinsic motivation levels were similar, with no statistically significant differences observed in either group. This outcome suggests that the type of feedback provided did not exert a substantial influence on the intrinsic motivation levels of the participants. Similarly, the findings concerning extrinsic motivation levels indicated that these levels were consistent across both groups and remained stable over time. The persistence of extrinsic motivation levels across different types of feedback and throughout the duration of the intervention suggests that external rewards or pressures either did not vary significantly during the study or that their impact was relatively minimal. The research findings on amotivation reveal a nuanced portrait, indicating a statistically significant increase in amotivation levels within the automated feedback group. In contrast, these levels remained constant within the teacher feedback group. This increase in amotivation may be attributed to a number of factors, including the relatively lower mean in the initial writing scores of the participants in the automated feedback group or the nature of the corrections provided through Paperpal, which were perceived as somewhat discouraging by some participants. It is also possible that the feedback provided through automated tools may have been somewhat disconnected from the instructional context of the courses, although the feedback prompt generated for ChatGPT was created in consideration of the rubric used by the instructor for the assignments, along with the types of the assignments.

The qualitative findings, derived through an analysis of interview data conducted with 6 volunteering participants from each group provide significant insights into the effective impact of writing feedback on L2 writing motivation. The results of the thematic analysis on the interview transcripts reveal a distinct level of differences between the teacher and automated feedback groups which highlight various factors that affect the writing motivation and anxiety of the participating learners.

Respondents in the control group often stated that feedback reception often led to heightened anxiety and a somewhat diminished sense of self-confidence particularly in the presence of ambiguity regarding which aspects of their writing should be improved and how they can be corrected. The fear of disappointing the teachers or one's self expectations also appear to be a source of anxiety within the teacher feedback group. This was also supported by those who made references to the issue of timely feedback, stating that they had either forgotten or have become disconnected from the writing assignment that they had submitted due to the late feedback that they had received, which resulted in either a lower level of willingness to work on improving the paragraphs or in more extreme cases, not submitting a second draft at all. While the workload on the teacher is also recognized by several respondents in both groups, the outcomes of having to provide comprehensive feedback to yeah substantial number of learners likely resulted in feedback that was recognized as ambiguous, undetailed or late by some of the respondents in the teacher feedback group.

On the other hand, the experimental group, which received automated feedback reported a different trajectory regarding their anxiety levels. While the existence of writing anxiety was acknowledged by the majority of the participants in the automated feedback group, frequently related to mistakes and a feeling of inadequacy, it was also frequently reported that true iterative, presumably non-judgmental, friendly and praise including feedback, their anxiety was decreased over time. Many in this group reported that through the recognition and correction of mistakes with the help of the feedback and received allowed their anxiety levels to decrease and eventually transform into motivation. This may suggest that the quick, detailed and friendly features of the automated feedback provided aided the students with the reduction of negative emotional reactions, supporting the positive developmental trajectory in their writing skills.

In improving the motivation levels of the participants, significant factors reported were positive reinforcement, personalization, constructive feedback, and recognition. A notable increase in motivation was reported by those in the automated feedback group, largely due to positive reinforcement and the personalized field of the comments provided through the genAI, which were constructed to include the combination of praise and criticism. Several times, participants in the automated feedback group praised the motivating effects of positive comments that ChatGPT was prompted to provide at the end of the written feedback. Feedback that included some trays was also viewed to be constructive, likely due to the fact that criticisms on one's writing demotivated the

participants. Those in the teacher feedback group made frequent statements that their motivation levels would increase through receiving praise regarding strong aspects of their paragraphs as well as criticism on mistakes. Suggestions on how to improve mistakes but also favored by those in the automated feedback group, while the lack thereof was criticized by some in the teacher feedback group while some also stated that they would enjoy receiving such suggestions. Feedback that conveyed a sense of audience and engagement from the provider such as comments that show the feedback provider relates to the story or discussion presented by the writers were also found to be highly valued, which was noted to provide encouragement and made the feedback process more satisfying for the recipients.

In terms of interest and extrinsic motivation levels, the teacher feedback group often cited the interpersonal interactions between the teacher and the student there's more to eating factors, while you automated feedback group emphasized personal growth and the improvement of writing skills to be what motivated them to write.

The participants in the automated feedback group were not informed that their feedback originated from ChatGPT. However, one participant, B3M, distinguished by frequent engagement and literacy of genAI tools, spotted AI-like patterns in the feedback, identifying its origin before disclosure. This respondent later remarked that the comments “felt repetitive” after receiving multiple instances of feedback. Despite this, he expressed surprise upon learning that the feedback was generated by ChatGPT. In a related observation within the teacher feedback group, multiple participants utilized ChatGPT to refine and clarify the feedback provided by human teachers, mirroring the practices observed in the automated feedback group. The findings of the present study on learners’ ability to differentiate between human and AI-generated texts reveal similar findings with those present in the literature, such as those of Hostetter et al. (2023). Consequently, it may be proposed that by meticulously designing and rigorously testing prompts, educators can offer learners a supportive, authentic, and advantageous source of feedback.

These insights underscore the significant role that genAI tools can play in educational settings, particularly in writing instruction. By integrating such tools, students can access an alternative source of feedback, enhancing their learning experience, while teachers can alleviate some of their workload through AI assistance. Moreover, the introduction and standardization of AI tools in educational contexts can serve dual purposes: enhancing instructional strategies and preventing unethical AI use through the development of informed and responsible habits among students. Teachers, too, stand to benefit immensely from the integration of AI in various educational practices, not only in providing feedback but also in other teaching-related tasks. This approach would not only familiarize both students and teachers with AI-generated content but would also promote the development of critical skills necessary to identify and utilize such content effectively for teachers.

Participants' views about artificial intelligence and ChatGPT varied, with some of the participants stating that feedback generated by such algorithms might be repetitive and generic in comparison to teachers' feedback, which was praised to be include personal touch. The authoritative role of the teacher is also mentioned within this context by multiple participants, particularly in the teacher feedback group. The interpersonal relationships and the assurance derived from engaging with a familiar, visible, and knowledgeable figure was also highlighted. Participants in the automated feedback group, such as B3M, expressed concerns regarding the limitations inherent in artificial intelligence tools. These limitations, including a perceived lack of genuine commentary and personal insight, rendered the feedback from such tools less satisfactory when compared to that from human instructors. The utilization of ChatGPT was viewed with considerable skepticism by the students, who equated its use with academic dishonesty. This perception of ChatGPT as a facilitator of cheating, rather than as an educational resource, was widespread among many respondents from both groups. These respondents felt as though they were under suspicion merely for being asked about their awareness or use of ChatGPT, likely due to said perception. It has been observed that the practice of withholding the information that feedback provided to the experimental group was generated by ChatGPT has prevented the alteration regarding the feedback they had received. It is also understood from the findings of the qualitative data that learners' perceived usefulness, attitude and computer anxiety would have a larger negative effect on their writing motivation and could further increase their anxiety, as highlighted by Li et al. (2019).

Notably, one participant, A5F, remarked that she would feel less embarrassed about her errors if the feedback were provided from ChatGPT, rather than the teacher. This aspect of automated feedback, also highlighted by previous studies (e.g. Aljanabi et al., 2023; Xiao & Zhi, 2023), could be exceptionally beneficial for students who experience lower confidence and heightened anxiety about writing, particularly fears of judgment, mockery, or criticism. Despite a general preference for human-generated and face-to-face feedback, especially among those in the teacher feedback group, the revelation that the feedback they had received was largely generated by ChatGPT and Paperpal led to a surprisingly positive response from the majority of participants in the automated feedback group. They appreciated the personalized, detailed, and human-like quality of the feedback. Such perspectives indicate a nuanced reception to automated feedback systems, suggesting that under certain conditions, such tools can emulate the qualities of human interaction and potentially enhance the learning experience.

The selection of writing topics, subjects, and genres has been identified as a critical factor influencing students' willingness to engage in writing tasks, as noted by participants in both the control and experimental groups. The majority of respondents highlighted that topics misaligned with their interests or knowledge were a significant source of anxiety. This anxiety stemmed from the additional time required to plan and conceptualize the text, particularly under the time constraints of exams or in-class assignments. Writing topics aligned with students' interests and knowledge were

cited as pivotal in enhancing motivation and reducing anxiety by respondents such as A6M, B4F, and B5F. Furthermore, participants like B2F and B3M suggested that allowing students to select their own writing topics, rather than assigning predetermined ones, could markedly improve both motivation and overall writing performance. This sentiment was supported by other participants, including B4F and B5F, who noted the positive impact of engaging topics on their motivation and anxiety levels. Similarly, the study by Sağlamel and Kayaoğlu (2015) highlighted the impact of topic selection on writing motivation and enjoyment. It was observed that topics aligning with students' interests tend to enhance their enthusiasm for writing, whereas topics outside their comfort zone in terms of interest and knowledge can induce anxiety. The results of the current investigation exhibit substantial resemblance with the findings documented in the research conducted by Sağlamel and Kayaoğlu utilizing comparable samples from different periods. The findings with regards to writing topics were also similar to those of the recent study by Eryılmaz (2023), which has also emphasized that learners prefer more flexible or free writing topics. Providing students with the freedom to choose their subjects of writing appears to remain a beneficial strategy in promoting better writing outcomes and enhancing academic performance.

Findings regarding preferred method of feedback reveal that participants generally displayed a strong performance for face-to-face feedback from instructors in class over alternated or online feedback. This sentiment was rooted in the release that feedback provided in real time through a direct interaction is more detailed and memorable enhancing the understanding and retention of the feedback. B4F, for example noted that face to face feedback was more effective due to the role of the human element. The role of positive feedback and praise was also frequently referred to throughout the interviews, participants stated that they value feedback that praised and acknowledged what was done well in the writing as well as pointing out areas for improvement. This was a feature of the automated feedback that was praised and appreciated by those in the experimental group. While those in the teacher feedback group also voiced their perceived value of positive feedback, they stated that they did not receive enough positive statements regarding their writing. This difference in the perception of the existence and recognition of positive feedback may be caused by inherent differences between the methods of AI-assisted and teacher feedback, where ChatGPT, as it was prompted, made friendly comments and positive suggestions regarding the writing off participants wild teacher feedback was relatively more focused on correcting any improving the students’ mistakes.

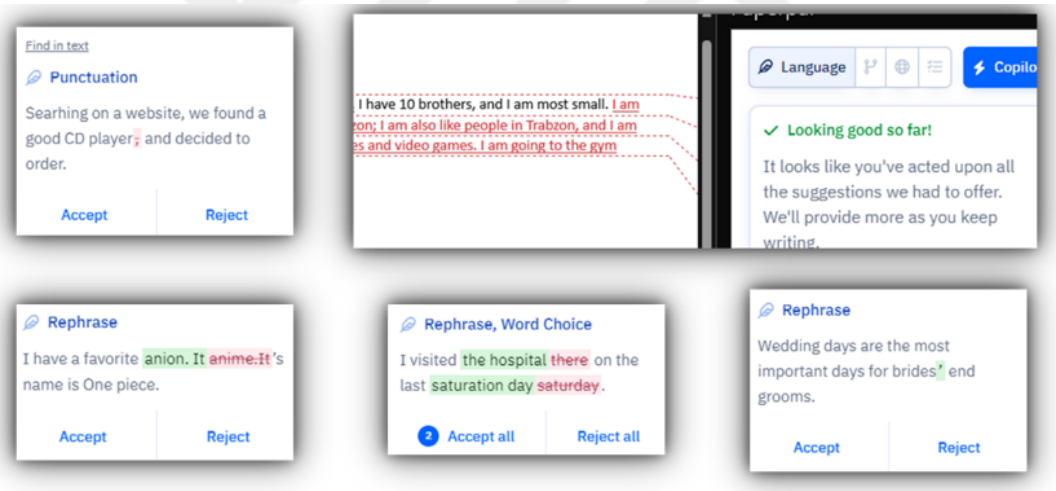
Feedback on content, provided by ChatGPT in the experiment group and by the instructor in the control group, received numerous references throughout the interviews. Suggestions and directions regarding the enhancement of argument strength, sentence organization, and overall paragraph content are perceived as crucial by participants in both the experimental and control groups. Participants in the experimental group, who received automated feedback, valued the tailored content-related comments and suggestions designed to aid in recognizing and ameliorating issues

related to organization, argumentation, and the fluidity of their writing. Conversely, participants in the control group reported dissatisfaction due to the absence of such detailed feedback on paragraph content, highlighting a potential area for improvement to create an educational environment that maximizes each student's potential by focusing on both content and form. These observations underscore the pivotal role of substantive feedback, noting that participants who received less feedback on content voiced their dissatisfaction, whereas those who received automated feedback were satisfied. The observations regarding the predominant preferences for feedback methods and content among Turkish EFL students align with the results presented by Tasdemir and Yalcin Arslan (2018), who similarly highlighted that students prefer teacher feedback that is both explicit and supplemented with detailed explanations. One other concern voiced by those within the teacher feedback group was that while they were provided with critical information regarding their writing skills, mistakes and lacks that could be improved, they were not notified how to improve specific sentences, phrases, and words within the specific paragraph they had submitted. Such a complaint was not made by any of the participants in the automated feedback group apart from two respondents, B1M and B6F, who only highlighted a singular instance where corrections made by Paperpal we are not understood correctly by the respondents. On learners' renposes to AWE, the findings of the present study bear similarities to those present in the literature that have investigated student satisfaction (Wei, 2015; Zhou, 2013). The responses regarding the significance of explicit and detailed feedback on learners' writing bear resemblance to those documented by Uzun and Köksal (2020) in their investigation of Turkish EFL students, which demonstrated a preference among the students for explicit feedback delivered in a face-to-face manner. Through the implementation of automated tools and genAI, issues with feedback on explicitness, content, and suggestions that support corrections could also be solved.

Throughout the study, observations were made concerning the effectiveness of automated feedback tools, including ChatGPT and Paperpal. ChatGPT often delivered comprehensive and detailed feedback when prompted. However, its performance was inconsistent, with some instances marked by the provision of shorter texts and lists rather than full paragraphs. This inconsistency underscores the critical role of the instructor in moderating the use of such feedback mechanisms. A notable limitation of ChatGPT's feedback was its inability to recognize context that would typically be apparent to an educator, such as errors in student writing related to personal context or idiomatic expressions directly translated from Turkish, which lose their intended meaning. Despite ChatGPT's extensive knowledge base, it lacked the capability to tailor its output to individual students without specific prompts. Previous studies, which informed the planning and thorough evaluation of the methodologies employed in the current study, have consistently reported similar findings regarding inaccuracies and originality (Alkaissi & McFarlane, 2023). Paperpal, utilized primarily for formative feedback, exhibited deficiencies in identifying and correcting errors common among Turkish EFL students, including spelling and grammatical errors. This limitation was compounded by the requirement for digital transcription of handwritten assignments, introducing further potential for

error. As Paperpal was initially designed to provide form-related feedback on academic writing, feedback it provided may have been slightly over the level of the participating students, who often made mistakes in word choice, spelling, and grammar, which may have gone under the radar of Paperpal. At multiple instances, corrections of Paperpal were either rejected or replaced by the researcher to avoid confusion on the part of the receiving students. Despite occasional imperfections, Paperpal enables the provision of feedback on form within minutes for assignments across an entire classroom, a task that would otherwise require considerable time and effort from the instructor. The data collected from interviews indicates minimal references to dissatisfaction among participants in the experimental group with the feedback on form provided by Paperpal, except for a few instances of implicit corrections. Given these factors, it can be asserted that while the implementation of Paperpal in providing feedback to preparatory grade EFL students is efficient, it necessitates the incorporation of teacher insights and filtering due to its tendency for academic texts, as suggested in the conclsinons of Grimes and Warschauer (2010).

Figure 6: Errors in Punctuation, Rephrasing, and Word Choice Corrections Detected on Paperpal



Throughout the intervention, observations, as well as statements from participants B1M and B6F, indicate that alternative methods may be considered for providing feedback on form in the context of introductory writing courses at the EFL preparatory level. The conclusion drawn from the use of both tools under study highlights that they frequently required human intervention or modifications to the prompts to achieve effective outcomes. The data pertaining to detailed suggestions and written feedback on content reveal a promising potential for the incorporation of genAI tools into the practices of writing feedback. Utilizing such tools could enable educators to provide comprehensive, friendly, and timely feedback, thereby reducing the requisite time and effort to deliver detailed feedback on each student's assignments. As suggested by Vojak et al. (2011), the social aspect of writing instruction and the critical role of teacher-student interactions through assignments and feedback should not be ignored in writing instruction, regardless of the success of

any implemented technology that can provide writing feedback. Therefore, it is crucial to acknowledge the substantial differences in time consumption and effort between manual teacher feedback and automated feedback. This distinction served as the primary reason for selecting ChatGPT as an alternative or supplementary feedback mechanism in this research.

The findings of the present study offer a range of implications for teaching practices, educational policymaking, and their effects on motivation, anxiety, and engagement. The notable reduction in cognitive anxiety across both groups suggests that the incorporation of AWE and GenAI tools for constructive feedback in writing instruction holds substantial promise in enhancing learning experiences and maintaining relatively low anxiety levels, while positively influencing learners' willingness to write. Despite certain drawbacks, such as a slight increase in amotivation, the intervention demonstrated effectiveness comparable to teacher feedback, as evidenced by significant differences favouring the automated feedback group. This indicates that the use of such tools in writing instruction warrants further investigation and experimentation, particularly concerning long-term writing performance. Additionally, the highly praised positive impact of constructive and personalized comments generated by tools like ChatGPT suggests that various educational levels and branches could effectively integrate automated feedback tools in their writing instruction classrooms, alongside the diverse utilities that LLMs offer. The ability of these tools to alleviate teachers' workload and save critical time, which can be redirected to other classroom issues, is accompanies by the timeliness and detail of feedback provided by ChatGPT and Paperpal. Nevertheless, ongoing research into the effects on writing performance and secondary studies on the longitudinal impacts are necessary to ensure that the corrections and suggestions provided by these tools positively influence learners' writing skills and academic performance in writing classrooms.

The present study was undertaken with the aim of understanding the differences between teacher and automated feedback to English writing assignments and to assess differences between the anxiety and motivation levels of learners that received said types of feedback. The main focus of the study was two explore the potentials of automated feedback methods and genAI tools, particularly ChatGPT, a large language model, which was used to provide feedback on content, supported with Paperpal, which provided feedback on form. From the observations and analysis of the procedure, the following suggestions can be made for further inquiry:

1. Conducting a longitudinal study exploring the long-term effects of automated feedback tools and genAI could prove to be useful, as the long-term effects of the feedback provided by ChatGPT are currently unknown.
2. The present study focused mainly on anxiety and motivation in relation to writing assignments. For further studies investigating LLMs, an analysis of the development of learners' writing performance could also be undertaken to assess differences in the various stages of writing development.

3. The present study was aimed at comparing the differences between teacher and automated feedback. However, as the results indicate, a methodology where teacher and automated feedback are used in tandem could also be explored.
4. LLMs could be experimented with in terms of how they can interpret and provide feedback with the gradual development in students' assignments in consideration.

REFERENCES

Acat, Mehmet Bahaddin and Demiral, Seyfettin, (2002), "Sources of Motivation in Learning Foreign Language in Turkey", **Educational Administration in Theory and Practice**, 31(31), 312-329.

Agarwal, Ritu and Prasad, Jayesh, (1999), "Are Individual Differences Germane to the Acceptance of new Information Technologies?", **Decision sciences**, 30(2), 361-391.

Balıkcıoğlu Akkuş, Gülin and Altay, İsmail Fırat (2023), "Instructor's Written Feedback in Emergency Remote Teaching: EFL Learners' Perspectives", **Bartın University Journal of Faculty of Education**, 12(1), 47-56.

Aktay, Sayım and Gök, Seçkin (2023), "ChatGPT in Education", **Türk Akademik Yayınlar Dergisi (TAY Journal)**, 7(2), 378-406.

Akyuz, Yaşar (2020), "Effects of Intelligent Tutoring Systems (ITS) on Personalized Learning (PL)", **Creative Education**, 11, 953-978.

Alafnan, Mohammad Awad et al. (2023), "Chatgpt as an Educational Tool: Opportunities, Challenges, and Recommendations for Communication, Business Writing, and Composition Courses", **Journal of Artificial Intelligence and Technology,** *3*(2), 60-68.

Alaraj, M. Mamoon (2016), "EFL acquisition external problems and difficulties: An Investigative Interview to Examine Saudi Students' Thoughts and Feelings", **Journal of Language Teaching and Research**, 7(3), 484-490.

Alharbi, Wael (2023), "AI in the Foreign Language Classroom: A Pedagogical Overview of Automated Writing Assistance Tools", **Education Research International**, 2023, 1-15.

Al Hosni, Samira (2014), "Speaking Difficulties Encountered by Young EFL Learners", **International Journal on Studies in English Language and Literature (IJSELL),** 2(6), 22-30.

Ali, Jamal Kaid Mohammed, et al. (2023), "Impact of ChatGPT on Learning Motivation: Teachers and Students' Voices", **Journal of English Studies in Arabia Felix**, 2(1), 41-49.

Aljanabi, Mohammed et al. (2023), "ChatGPT: Open Possibilities", **Iraqi Journal for Computer Science and Mathematics**, 4(1), 62-64.

Alkaissi, Hussam and McFarlane, Sammy I (2023), "Artificial hallucinations in ChatGPT: implications in scientific writing", **Cureus**, 15(2).

Alshami, Ahmad et al. (2023), "Harnessing the Power of ChatGPT for Automating Systematic Review Process: Methodology, Case Study, Limitations, and Future Directions", **Systems**, 11(7), 351-388.

Amalia, Hilma et al. (2021), "Teaching Writing to Junior High School Students: A Focus on Challenges and Solutions". **Journal of Language and Linguistic Studies**, *17*(S2), 794-810.

Andrade, Maureen Snow and Evans, Norman W (2012), **Principles and Practices for Response in Second Language Writing: Developing Self-Regulated Learners**. Routledge, London.

Snow, Jacob (2018, July 26), "Amazon's facial recognition tool misidentified 28 members of Congress as people arrested for crimes", **ACLU**. https://www.aclu.org/news/privacy-technology/amazons-face-recognition-falsely-matched-28 (10.11.2023).

Arulkumaran, Kai et al. (2017), "Deep Reinforcement Rearning: A Brief Survey", **IEEE Signal Processing Magazine**, 34(6), 26-38.

Aslan, Imran et al (2020), "Exploring perceived stress among students in Turkey during the COVID-19 pandemic", **International Journal of Environmental Research and Public Health**, 17(23), 8961.

Astriningsih, Nurhapsari and Mbato, Concilianus Laos (2019), "Motivation to learn English: Why Indonesian Adult Learners Join a Community of Interest", **Anima Indonesian Psychological Journal**, 34(2), 65-75.

Attali, Yigal (2013), "Validity and Reliability of Automated Essay Scoring", In **M. D. Shermis & J. Burstein (Eds.), Handbook of automated essay evaluation** (181-198), Routledge, New York.

Aydın, Ömer and Karaarslan, Enis (2023), "Is ChatGPT leading generative AI? What is Beyond Expectations?", **Academic Platform Journal of Engineering and Smart Systems**, 11(3), 118-134.

Aydın, Selami (2008), "An Investigation on the Language Anxiety and Fear of Negative Evaluation Among Turkish EFL Learners", **Asian EFL Journal**, **Teaching Articles**, 30(1), 421-444.

Aydın, Selami (2012), "The Effects of Young EFL Learners' Perceptions of Tests on Test Anxiety", **Education**, 40(2), 189-204.

Baidoo-Anu, David and Ansah, Leticia Owusu (2023), "Education in The Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of Chatgpt in Promoting Teaching and Learning", **Journal of AI**, 7(1), 52-62.

Baloğlu, Gürol and Çakalı, Kaan Ramazan (2023), "Is Artificial Intelligence a New Threat to the Academic Ethics?: Enron Scandal Revisited by ChatGPT", **İşletme**, 4(1), 143-165.

Bao, Minhui (2019), "Can Home use of Speech-Enabled Artificial Intelligence Mitigate Foreign Language Anxiety–Investigation of a Concept", **Arab World English Journal (AWEJ) Special Issue on CALL**, (5), 28-40.

Barrot, Jessie S. (2021), "Using automated written corrective feedback in the writing classrooms: Effects on L2 writing accuracy", **Computer Assisted Language Learning**, 4(36), 584-607.

Bawarith, Razan et al. (2017), "E-exam Cheating Detection System", **International Journal of Advanced Computer Science and Applications**, 8(4), 176-184.

Baydere, Hilal Öztürk (2023), "Rethinking Translator Competence in the Machine Translation Era", **TransLogos Translation Studies Journal**, 6(1), 45-74.

Beck, Joseph et al. (1996), "Applications of AI in Education", **XRDS: Crossroads, The ACM Magazine for Students**, 3(1), 11-15.

Beikian, Ali and Ganji, Mansoor (2022), "Efficacy of Instatext for Improving Persian-English Freelance Translators' Language Quality: From Perception to Practice", **Journal of Foreign Language Teaching and Translation Studies**, 7(4), 59-86.

Bektaş-Çetinkaya, Yeşim & Oruç, Nesrin (2010), "Turkish Students' Motivation to Learn English at Public and Private Universities", **Procedia-Social and Behavioral Sciences**, 2(2), 4662-4666.

Benson, Phil (2013), **Teaching and researching: Autonomy in language learning**, 2nd Ed., Routledge, Harlow.

Berkey, Dennis and Halfond, Jay (2015), "Cheating, Student Authentication and Proctoring in Online Programs", **New England Board of Higher Education**. https://nebhe.org/journal/cheating-student-authentication-and-proctoring-in-online-programs/ (10.12.2023).

Betz, Nancy E. (1978), "Prevalence, Distribution, and Correlates of Math Anxiety in College Students", **Journal of Counseling Psychology**, 25, 441–448.

Beuving, Joost and de Vries, Geert (2015), **Doing Qualitative Research: The Craft of Naturalistic Inquiry**, Amsterdam University Press, Amsterdam.

Biddle, Sam (2022, December 9), "The internet's new favorite AI proposes torturing Iranians and surveilling mosques", **The Intercept**. https://theintercept.com/2022/12/08/openai-chatgpt-ai-bias-ethics/ (10.12.2023).

Bilen, Eren and Matros, Alexander (2021), "Online Cheating Amid COVID-19", **Journal of Economic Behavior & Organization**, 182, 196-211.

Birjandi, Parviz and Hadidi Tamjid, Nasrin (2012), "The Role of Self-, Peer and Teacher Assessment in Promoting Iranian EFL Learners' Writing Performance", **Assessment & Evaluation in Higher Education**, 37(5), 513-533.

Biswas, Som (2023), "ChatGPT and the future of medical writing", **Radiology**, 307(2), e223312.

Bitchener, John et al. (2005), "The Effect of Different Types of Corrective Feedback on ESL Student Writing", **Journal of Second Language Writing**, 14(3), 191-205.

Bozkurt, Aras and Sharma, Ramesh C. (2023), "Generative AI and Prompt Engineering: The Art of Whispering to Let the Genie tut of the Algorithmic World", **Asian Journal of Distance Education**, 18(2).

Braun, Virginia and Clarke, Victoria (2006), "Using Thematic Analysis in Psychology", **Qualitative Research in Psychology**, *3*(2), 77-101.

Brown, James Dean (1998), **New Ways of Classroom Assessment. New Ways in TESOL Series II. Innovative Classroom Techniques**. TESOL, Alexandria.

Bruffaerts, Ronny et al. (2018), "Mental Health Problems in College Freshmen: Prevalence and Academic Functioning", **Journal of Affective Disorders**, 225, 97-103.

Bruno, Inês and Santos, Leonor (2010), "Written Comments as a Form of Feedback", **Studies in Educational Evaluation**, *36*(3), 111-120.

Buriak, Jillian M., et al. (2023), "Best Practices for Using AI When Writing Scientific Manuscripts: Caution, Care, and Consideration: Creative Science Depends on it", **ACS Nano**, 17(5), 4091-4093.

Burstein, Jill et al. (2004), "Automated Essay Evaluation: The Criterion Online Writing Service", **AI Magazine**, 25(3), 27-27.

Burstein, Jill et al. (2016), "Informing Automated Writing Evaluation Using the Lens of Genre: Two Studies", **Calico Journal**, 33(1), 117-141.

Burstein, Jill et al. (2020), "Expanding Automated Writing Evaluation", Yan et al. (Eds.) **In Handbook of Automated Scoring** (329-346), Chapman and Hall/CRC, Florida.

Cahill, Aoife and Evanini, Keelan (2020), "Natural language processing for writing and speaking", Yan et al. (Eds.), In **Handbook of Automated Scoring** (69-92), Chapman and Hall/CRC, Florida.

Calma, Angelito et al. (2022), "Grammarly: An Instructional Intervention for Writing Enhancement In Management Education", **The International Journal of Management Education**, 20(3), 100704.

Campbell, Donald T., and Stanley, Julian C. (2015), **Experimental and Quasi-Experimental Designs for Research.** Ravenio Books. London.

Cao, Lien Thi Xuan et al. (2023), "Technology-Assisted Task-Based Second Language Assessment for Learning". Bui et al. (Eds.) **In Innovations and Applications of Technology in Language Education**, 130-146, Auerbach Publications, New York.

Carson, Joan G and Nelson, Gayle L. (1994), "Writing Groups: Cross-Cultural Issues", **Journal of Second Language Writing**, 3(1), 17-30.

Carvalhais, Lenia et al. (2020), "Effects of The Portuguese Graphogame on Reading, Spelling, and Phonological Awareness in Second Graders Struggling to Read", **Journal of Writing Research**, 12(1), 9–34.

Celik, Ismail et al. (2022), "The Promises and Challenges of Artificial Intelligence for Teachers: A Systematic Review of Research", **TechTrends**, 66(4), 616-630.

Chaka, Chaka (2023), "Detecting AI Content in Responses Generated by ChatGPT, Youchat, and Chatsonic: The Case of Five AI Content Detection Tools", **Journal of Applied Learning and Teaching**, 6(2), 94-104.

Chapelle, Carol A. (2010), "The Spread of Computer-Assisted Language Learning", **Language Teaching**, 43(1), 66-74.

Chapelle, Coral A. et al. (2015), "Validity Arguments for Diagnostic Assessment Using Automated Writing Evaluation", **LanguageTesting**, 32(3), 385-405.

Chen, Chi-Fen Emily and Cheng, Wei-Yuan Eugene Cheng Cheng (2008), "Beyond the Design of Automated Writing Evaluation: Pedagogical Practices and Perceived Learning Effectiveness in EFL Writing Classes", **Language Learning & Technology,** 12(2), 94–112.

Chen, Tzeng-Ji (2023), "Chatgpt and Other Artificial Intelligence Applications Speed Up Scientific Writing", **Journal of the Chinese Medical Association**, 86(4), 351-353.

Cheng, Gary (2017), "The Impact of Online Automated Feedback on Students' Reflective Journal Writing in an EFL Course", **The Internet and Higher Education**, 34, 18-27.

Cheng, Yuh-Show (2002), "Factors Associated with Foreign Language Writing Anxiety", **Foreign Language Annals**, 35(6), 647-656.

_____ (2004), "A Measure of Second Language Writing Anxiety: Scale Development and Preliminary Validation", **Journal of Second Language Writing**, 13(4), 313-335.

Chew, Chiou Sheng et al. (2020), "Enhancing Summary Writing of ESL Learners Via a Theory-Based Online Tool: System Development and Evaluation", **Journal of Educational Computing Research**, 58(2), 398-432.

Chocarro, Raquel et al. (2023), "Teachers' Attitudes Towards Chatbots in Education: A Technology Acceptance Model Approach Considering the Effect of Social Language, Bot Proactiveness, and Users' Characteristics", **Educational Studies**, 49(2), 295-313.

Choi, Inn-Chull (2011), "Efficacy of an Automatic Essay Scoring-Based English Writing In-Service Training Program", **Multimedia Assisted Language Learning**, 14(2), 61-89.

Choi, Seongyune et al. (2023), “Influence of Pedagogical Beliefs and Perceived Trust on Teachers’ Acceptance of Educational Artificial Intelligence Tools”, **International Journal of Human–Computer Interaction**, 39(4), 910-922.

Chou, Huey-Nah Cindy et al. (2016), “My Access and Writing Error Corrections of EFL College Pre-Intermediate Students”, **International Journal of Education**, 8(1), 144-161.

Christensen, Clayton M. et al. (2011), **Disruptive Innovation: The Christensen Collection (The Innovator's Dilemma, The Innovator's Solution, The Innovator's DNA, and Harvard Business Review article” How Will You Measure Your Life?”)(4 Items),** Harvard Business Press, Boston.

Christensen, Larry. B., et al. (2011), **Research Methods, Design, and Analysis**. 12th Edition, Pearson, Alabama.

Chui, Michael et al. (2018), “Notes from the AI frontier: Applications and Value of Deep Learning”, **McKinsey Global Institute**.

Chun, Dorothy et al. (2016), “Technology in Language Use, Language Teaching, and Language Learning”, **The Modern Language Journal**, 100(S1), 64-80.

Clerwall, C. (2014), Enter the Robot Journalist: Users’ Perceptions of Automated Content. **Journalism Practice,** 8(5), 519–531.

Cohen, Andrew D. and Brooks-Carson, Amanda (2001), “Research on Direct Versus Translated Writing: Students' Strategies and Their Results”, **The Modern Language Journal**, 85(2), 169-188.

Cohen, Yoav et al. (2018), “Validating Human and Automated Scoring of Essays Against “True” Scores”, **Applied Measurement in Education**, 31(3), 241-250.

Comas-Forgas, Rubén et al. (2021), “Exam Cheating and Academic Integrity Breaches During the COVID-19 Pandemic: An Analysis of Internet Search Activity in Spain”, **Heliyon**, 7(10).

Cornu, Bernard (1995), “New Technologies: Integration into Education”, Deryn Watson David Tinsley (Eds.) In **Integrating Information Technology into Education**, (3-11) Springer, New York.

Cotos, Elena (2011), “Potential of Automated Writing Evaluation Feedback”, **Calico Journal**, 28(2), 420-459.

_____ (2014), **Genre-based Automated Writing Evaluation for L2 Research Writing: From Design to Evaluation and Enhancement**. Springer, New York.

Crawford, Joseph et al. (2023), “Leadership Is Neededfor Ethical ChatGPT: Character, Assessment, and Learning Using Artificial Intelligence (AI)”, **Journal of University Teaching & Learning Practice**, 20(3), 02.

Creswell, John W. and Creswell, J. David (2017), **Research Design: Qualitative, Quantitative, and Mixed Methods Approaches**. (4th Ed.) Sage publications, Los Angeles.

Creswell, John W. and Poth, Cheryl N. (2016), **Qualitative Inquiry and Research Design: Choosing Among Five Approaches**. Sage publications, Los Angeles.

Crossley, Scott A. et al. (2014), "Linguistic Microfeatures to Predict L2 Writing Proficiency: A Case Study in Automated Writing Evaluation", **Journal of Writing Assessment**, 7(1).

Daly, John A. (1979), "Writing Apprehension in the Classroom: Teacher Role Expectancies of the Apprehensive Writer", **Research in the Teaching of English**, 13(1), 37-44.

Daly, John A. and Miller, Michael D. (1975), "The Empirical Development of an Instrument of Writing Apprehension", **Research in the Teaching of English**, 9, 242-249.

Dar, Muhammad Fareed and Khan, Imran (2015), "Writing Anxiety Among Public and Private Sectors Pakistani Undergraduate University Students", **Pakistan Journal of Gender Studies**, 10(1), 157-172.

Davis, Fred D. (1985), **A technology Acceptance Model for Empirically Testing New End-user Information Systems: Theory and Results,** Unpublished Doctoral dissertation, Massachusetts Institute of Technology.

_____ (1989), "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology", **MIS Quarterly**, 13(3), 319-340.

Deane, Paul (2013), "On the Relation Between Automated Essay Scoring and Modern Views of the Writing Construct", **Assessing Writing**, 18(1), 7-24.

Deci, Edward L. (1972), "Intrinsic Motivation, Extrinsic Reinforcement, and Inequity", **Journal of Personality and Social Psychology**, 22(1), 113.

Dekhinet, Rayenne (2008), "Online Enhanced Corrective Feedback for ESL Learners in Higher Education", **Computer Assisted Language Learning**, 21(5), 409-425.

Demirdaş, Özkan and Bozdoğan, Derya (2013), "Foreign Language Anxiety and Performance of Language Learners in Preparatory Classes", **Turkish Journal of Education**, 2(3), 4-13.

Dewaele, Jean-Marc and Al-Saraj, Taghreed (2015), "Foreign Language Classroom Anxiety of Arab Learners of English: The Effect of Personality, Linguistic and Sociobiographical Variables", **Studies in Second Language Learning and Teaching**, 5(2), 205-228.

Dikli, Semire and Bleyle, Susan (2014), "Automated Essay Scoring Feedback for Second Language Writers: How Does it Compare to Instructor Feedback?", **Assessing Writing**, 22, 1-17.

Dönmez, İsmail et al. (2023), "Conducting Academic Researchwith The AI Interface ChatGPT: Challenges and Opportunities", **Journal of STEAM Education**, 6(2), 101-118.

Dörnyei, Zoltán (1994), "Motivation and Motivating in the Foreign Language Classroom", **The Modern Language Journal**, 78(3), 273-284.

Dörnyei, Zoltán (1998), "Motivation in Second and Foreign Language Learning", **Language Teaching**, 31(3), 117-135.

Dörnyei, Zoltán (2001), **Motivational Strategies in the Language Classroom** 10th Ed., Cambridge University Press, Cambridge.Dörnyei, Zoltán (2007), **Research Methods in Applied Linguistics**, Oxford University Press, Oxford.

Dudeney, Gavin and Hockly, Nicky (2012), "ICT in ELT: How Did We Get Here and Where are We Going?", **ELT Journal**, 66(4), 533-542.

Duong, Michael Tran et al. (2019), "Artificial Intelligence for Precision Education in Radiology", **The British Journal of Radiology**, 92(1103), 20190389.

Ekmekçi, Emrah (2018), "Exploring Turkish EFL students' Writing Anxiety", **The Reading Matrix: An International Online Journal**, 18(1), 158-175.

Elaldi, Senel (2016), "Foreign Language Anxiety of Students Studying English Language and Literature: A Sample from Turkey", **Educational Research and Reviews**, 11(6), 219-228.

El Ebyary, Khaled and Windeatt, Scott (2010), "The Impact of Computer-Based Feedback on Students' Written Work", **International Journal of English Studies**, 10(2), 121-142.

Ellis, Rod (2003), "Task-Based Language Learning and Teaching", **ELT Journal,** 59(1), 75-77.

Eryılmaz, Ali and Yeşilyurt, Yusuf Emre (2022), "How to Measure and Increase Second Language Writing Motivation", **LEARN Journal: Language Education and Acquisition Research Network**, *15*(2), 922-945.

Eryılmaz, Ramazan (2023), "Opinions of Students Learning Turkish as a Foreign Language on the Use of E-Portfolio in Writing Lessons", **International Journal of Education and Literacy Studies**, 11(4), 85-97.

Eysenbach, Gunther (2023), "The Role of Chatgpt, Generative Language Models, and Artificial Intelligence in Medical Education: A Conversation with Chatgpt and a Callfor Papers", **JMIR Medical Education**, 9(1), e46885.

Faghih, Behnam, et al. (2014), "User Interface Design for E-Learning Software". **The International Journal of Soft Computing and Software Engineering [JSCSE],** 3(3), 786-794.

Fallon, Virginia (2023, April 23), "Students Caught Cheating with ChatGPT Offered Amnesty for Confession", **Etico**. https://etico.iiep.unesco.org/en/students-caught-cheating-chatgpt-offered-amnesty-confession (12.10.2023).

Fares, Omar H. et al. (2022), "Utilization of Artificial Intelligence in The Banking Sector: A Systematic Literature Review", **Journal of Financial Services Marketing**, 28(4), 835-852.

Farrell, Thomas S. C. (2013), **Reflective Language Teaching: From Research to Practice**, A & C Black, London.

Fernández-Batanero, José-Maria et al. (2021), "Impact of Educational Technology on Teacher Stress and Anxiety: A Literature Review", **International Journal of Environmental Research and Public Health**, 18(2), 548.

Ferris, Dana (2007), "Preparing Teachers to Respond to Student Writing", **Journal of Second Language Writing**, 16(3), 165-193.

Ferris, Dana and Hedgcock, John (2005), **Teaching ESL composition: Purpose, Process, and Practice**. Mahwah, NJ: Erlbaum.

Fırat, Mehmet (2023), "Integrating AI Applications into Learning Management Systems to Enhance E-Learning", **Instructional Technology and Lifelong Learning**, 4(1), 1-14.

Fitria, Tira Nur (2021), "Grammarly as AI-powered English Writing Assistant: Students'Alternative for Writing English", **Metathesis: Journal of English Language, Literature, and Teaching**, 5(1), 65-78.

Foltz, Peter W et al. (1999), "The Intelligent Essay Assessor: Applications to Educational Technology", **Interactive Multimedia Electronic Journal of Computer-Enhanced Learning**, 1(2), 939-944.

Franks, Raphael and Maher, Rachel (2023, May 16), " "It's Russian Roulette" - Top Students Deny AI Cheating After Teachers Flunk Them", **NZ Herald**. https://www.nzherald.co.nz/nz/ai-and-chatgpt-cheating-allegations-at-cambridge-and-pukekohe-high-outrage-parents-of-top-students/A5TBH25N2NC4JG2ZOAUB7L3W6Y/ (10.12.2023)

Franzke, Marita et al. (2005), "Summary Street®: Computer Supportf or Comprehension and Writing"**, Journal of Educational Computing Research**, 33(1), 53-80.

Fu, Huixia and Liu, Xiaobin (2022), "EFL Learner Engagement in Automatic Written Evaluation", **Frontiers in Psychology**, 13**,** 871707.

Fu, Mengxia, and Li, Shaofeng, (2022), "The Effects of Immediate and Delayed Corrective Feedback On L2 Development", **Studies in Second Language Acquisition**, *44*(1), 2-34.

Fui-Hoon Nah, Fiona et al. (2023), "Generative AI and ChatGPT: Applications, Challenges, and AI-Human Collaboration", **Journal of Information Technology Case and Application Research**, 25(3), 277-304.

Fyfe, Paul (2022), "How to Cheat on Your Final Paper: Assigning AI for Student Writing", **AI & Society**, 38, 1395-1405.

Gangaiamaran, Ramya and Pasupathi, Madhumathi (2017), "Review on Use of Mobile Apps for Language Learning", **International Journal of Applied Engineering Research**, 12(21), 11242-11251.

Gao, Catherine et al. (2022), "Comparing Scientific Abstracts Generated by Chatgpt to Original Abstracts Using an Artificial Intelligence Output Detector, Plagiarism Detector, and Blinded Human Reviewers", **NPJ Dijit,** 6 (1).

Gao, Jianwu and Ma, Shuang (2019), "The Effect of Two Forms of Computer-Automated Metalinguistic Corrective Feedback", **Language Learning & Technology**, 23(2), 65–83.

Gardner, Robert (1985), **Social Psychologyand Second Language Learning: The Role of Attitudes and Motivation**. London: Edward Arnold.

_____ (2001), "Integrative Motivation and Second Language Acquisition", Zoltan Dornyei & Richard Schmidt (Eds.), In **Motivation and Second Language Acquisition** (1-19), **TESOL Quarterly,** Honolulu.

Gardner, Robert and Lambert, Wallace (1972), **Attitudes and Motivation in Second-Language Learning**. Rowley, MA: Newbury House.

Geertz, Clifford (1973), **The Interpretation of Cultures** Vol. 5019, Basic books, New York.

Ghasemi, Asghar and Zahediasl, Saleh (2012), "Normality Tests for Statistical Analysis: A Guide for Nonstatisticians." **International Journal of Endocrinology and Metabolism**, 10(2), 486–489.

Gibbs, Graham and Simpson, Claire (2004), "Conditions Under Which Assessment Supports Students' Learning", **Learning and Teaching in Higher Education**, 1(1), 3-31.

Giray, Louie (2023), "Prompt Engineering with ChatGPT: A Guide for Academic Writers", **Annals of Biomedical Engineering**, 51(12), 2629-2633.

Göçen, Ahmet and Aydemir, Fatih (2020), "Artificial Intelligence in Education and Schools", **Research on Education and Media**, 12(1), 13-21.

Godwin-Jones, Robert (2011), "Mobile Apps for Language Learning", **Language Learning & Technology**, 15(2), 2–11.

_____ (2023), "Emerging Spaces for Language Learning: AI Bots, Ambient Intelligence, and the Metaverse", **Language Learning & Technology**, 27(2), 6–27.

Gödde, Daniel et al. (2023), "A SWOT (Strengths, Weaknesses, Opportunities, and Threats) Analysis of ChatGPT in the Medical Literature: Concise Review", **Journal of Medical Internet Research**, 25, e49368.

Göksu, İdris et al. (2021), "Distance Education Amid a Pandemic: Which Psycho-Demographic Variables Affect Students in Higher Education?", **Journal of Computer Assisted Learning**, 37(6), 1539-1552.

Golda, T. Lilly (2015), "Exploring Reasons for Writing Anxiety: A Survey", **The Journal for English Language and Literary Studies**, 5(2), 40-44.

Goldberg, Amie et al. (2003), "The Effect of Computers on Student Writing: A Meta-Analysis of Studies From 1992 To 2002", **The Journal of Technology, Learning and Assessment**, 2(1), 1–51.

Goodman, Bryce and Flaxman, Seth (2017), "European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation." " **AI Magazine**, 38(3), 50-57.

Graham, Steve (2018), "Instructional feedback in writing", In Anastasiya Lipnevich & Jeffrey Smith (Eds.), The Cambridge handbook of instructional feedback (145–168), **Cambridge University Press**.

_____ (2022), Teaching writing in the digital age, Douglas Fisher (Ed.), "Educational psychology section; **Routledge Encyclopedia of Education (Online),** Taylor & Francis.

Graham, Steve et al. (2014), "Teaching Writing to Middle School Students: A National Survey", **Reading and Writing**, 27(6), 1015–1042.

Graham, Steve and Perin, Dolores (2007), "A Meta-Analysis of Writing Instruction for Adolescent Students", **Journal of Educational Psychology**, 99(3), 445-476.

Grimes, Douglas and Warschauer, Mark (2010), "Utility in a Fallible Tool: A Multi-Site Case Study of Automated Writing Evaluation", **The Journal of Technology, Learning and Assessment**, 8(6).

Guo, Biyang et al. (2023), "How Close is Chatgpt to Human Experts? Comparison Corpus, Evaluation, and Detection". https://dblp.org/rec/journals/corr/abs-2301-07597.html (10.12.2023)

Habiburrahim, Habib et al. (2021), "Scrutinizing Cheating Behavior Among EFL Students at Islamic Higher Education Institutions in Indonesia", **The Qualitative Report**, 26(3), 1033-1053.

Hagiwara, Masato (2021), **Real-World Natural Language Processing: Practical Applications with Deep Learning**. Manning Publications, New York.

Haleem, Abid et al. (2022), "An Era of Chatgpt as a Significant Futuristic Support Tool: A Study on Features, Abilities, and Challenges", **Benchcouncil Transactions on Benchmarks, Standards and Evaluations**, 2(4), 10008

Han, Turgay and Sarı, Elif (2022), "An Investigation on The Use of Automated Feedback in Turkish EFL Students' Writing Classes", **Computer Assisted Language Learning**, 37(4), 961-985.

Handini, Baiq Suprapti et al. (2022), "The Effect of Artificial Intelligent Technology Used (Duolingo Application) To Enhance English Learning", **ELLITE: J**ournal of English Language, Literature, and Teaching, 7(2), 86-94.

Hassanzadeh, Mohammad and Fotoohnejad, Samira (2021), "Implementing an Automated Feedback Program for a Foreign Language Writing Course: A Learner-Centric Study", **Journal of Computer Assisted Learning**, 37(5), 1494-1507.

Hattie, John and Timperley, Helen (2007), "The Power of Feedback", **Review of Educational Research**, 77(1), 81-112.

Hazaea, Abduljalil Nasr et al. (2021), "Emergency Remote English Language Teaching in the Arab League Countries: Challenges and Remedies", **Computer-Assisted Language Learning Electronic Journal**, 22(1), 201-222.

Hew, Khe Foon & Cheung, Wing Sum (2014), "Students' and Instructors' use of Massive Open Online Courses (MOOCs): Motivations and Challenges", **Educational Research Review**, 12, 45-58.

Hidayati, Tuti and Diana, Sari (2019), "Students' Motivation to Learn English Using Mobile Applications: The Case of Duolingo and Hello English", **JEELS (Journal of English Education and Linguistics Studies)**, 6(2), 189-213.

Higgins, Derrick and Heilman, Michael (2014), "Managing Wha We Can Measure: Quantifying the Susceptibility of Automated Scoring Systems to Gaming Behavior", **Educational Measurement: Issues and Practice**, 33(3), 36-46.

Hockly, Nickly (2019), "Automated Writing Evaluation", **ELT Journal**, 73(1), 82–88,

Holmes, Wayne et al. (2023), "Artificial Intelligence in Education", Christoph Stückelberg and Pavan Duggall (Eds.) **In Data Ethics: Building Trust (621-652)**, Globethics Publications, Switzerland.

Hong, Wilson Cheong Hin Hong (2023), "The Impact of Chatgpt on Foreign Language Teaching and Learning: Opportunities in Education and Research", **Journal of Educational Technology and Innovation**, 5(1), 37-45.

Horning, Alice (1987), **Teaching Writing as a Second Language,** SIU Press, USA.

Horwitz, Elaine (2001), "Language Anxiety and Achievement", **Annual Review of Applied Linguistics**, 21, 112-126.

Horwitz, Elaine et al. (1986), "Foreign Language Classroom Anxiety", **The Modern Language Journal**, 70(2), 125-132.

Hostetter, Autumn et al. (2023), "Student and Faculty Perceptions of Artificial Intelligence in Student Writing".

Hoyt, indsay Till et al. (2021), " "Constant Stress Has Become The New Normal": Stress And Anxiety Inequalities Among US College Students in the Time of COVID-19", **Journal of Adolescent Health**, 68(2), 270-276.

Hsu, Hui-Yin et al. (2008), "Using Audioblogs to Assist English-Language Learning: An Investigation into Student Perception", **Computer Assisted Language Learning**, 21(2), 181-198.

Hu, Yinghui and Zhao, Dan (2015), "A Comparative Study of Teacher Feedback and Automated Essay Scoring in College English Writing", **International Journal of Linguistics and Communication**, *3*(2), 82-97.

Huang, Jingshan and Tan, Ming (2023), "The Role of Chatgpt in Scientific Communication: Writing Better Scientific Review Articles", **American Journal of Cancer Research**, 13(4), 1148.

Huang, Min et al. (2020), "Emergency Remote Teaching of English as a Foreign Language During COVID-19: Perspectives from a University in China", **IJERI: International Journal of Educational Research and Innovation,** (15), 400-418.

Huang, Peiwan and Hwang, Yanling (2013), "An Exploration of EFL Learners' Anxiety and E-Learning Environments", **Journal of Language Teaching and Research**, 4(1), 27-35.

Huawei, Shi, and Aryadoust, Vahid (2023), "A Systematic Review of Automated Writing Evaluation Systems", **Education and Information Technologies**, 28(1), 771-795.

Huang, Shu and Renandya, Willy (2020), "Exploring the Integration of Automated Feedback Among Lower-Proficiency EFL Learners", **Innovation in Language Learning and Teaching**, 14(1), 15-26.

Huh, Myung-Hye (2001), "Translation Strategy in EFL Writing", **English Teaching**, 56(4), 75-92.

Huot, Brian (1990), "Reliability, Validity, And Holistic Scoring: What We Know and What We Need to Know", **College Composition and Communication**, 41(2), 201-213.

İçbay, Mehmet Ali (2005), "A SWOT Analysis on the University Entrance Examination in Turkey: A Case Study", **Mersin Üniversitesi Eğitim Fakültesi Dergisi**, 1(1), 126-140.

Ives, Bob and Cazan, Ana-Maria (2024), "Did The COVID-19 Pandemic Lead to an Increase in Academic Misconduct in Higher Education?", **Higher Education**, 87(1), 111-129.

Jaiswal, Akanksha and Arun, Joe (2021), "Potential of Artificial Intelligence for Transformation of the Education System in India", **International Journal of Education and Development Using Information and Communication Technology**, 17(1), 142-158.

Jang, So Young and & Choi, Shoi (2014), "Validating the Second Language Writing Anxiety Inventory for Korean College Students", **Advanced Science and Technology Letters**, 59, 81-84.

Javaid, Mohd et al. (2022), "Understanding the Potential Applications of Artificial Intelligence in Agriculture Sector", **Advanced Agrochem.** 2(1), 15-30.

Jereb, Eva et al. (2018), "Factors Influencing Plagiarism in Higher Education: A Comparison of German and Slovene Students". **PloS One**, 13(8), e0202252.

Ji, Ziwei et al. (2022), Yejin Bang, Andrea Madotto, and Pascale Fung. 2022. Survey of Hallucination in Natural Language Generation. **ACM Computing Surveys**, 55(12), 1-38.

Jiang, Liangjiang et al. (2020), "Second Language Writing Instructors' Feedback Practice in Response to Automated Writing Evaluation: A Sociocultural Perspective", **System**, 93, 102302.

Jourdan, Léane et al. (2023), "Text revision in scientific writing assistance: An overview", https://arxiv.org/pdf/2303.16726 (10.12.2023)

Jönsson, Anders et al. (2021), "Analytic or holistic? A Study About How to Increase the Agreement in Teachers' Grading", **Assessment in Education: Principles, Policy & Practice**, 28(3), 212-227.

Kalay, Serdal & Büyükkarcı, Kağan (2020), "English Language Teachers' Views on Teaching and Assessment of Writing Skills", **SDU International Journal of Educational Studies**, 7(2), 262-286.

Kane, Michael (2013), "Validating the Interpretations and Uses of Test Scores", **Journal of Educational Measurement**, 50(1), 1-73.

Kaplan, Randy et al. (1998), "Scoring Essays Automatically Using Surface Features", **ETS Research Report Series**, 1998(2), 1-12.

Kasneci, Enkelejda e al. (2023), "Chatgpt for Good? On Opportunities and Challenges of Large Language Models for Education", **Learning and Individual Differences**, 103, 102274.

Kay, Robin (2012), "Exploring The use of Video Podcasts in Education: A Comprehensive Review of the Literature", **Computers in Human Behavior**, 28(3), 820-831.

Kayaoğlu, Mustafa Naci and Sağlamel, Hasan (2013), "Students' Perceptions of Language Anxiety in Speaking Classes", **Journal of History Culture and Art Research**, 2(2), 142-160.

Kearns, Lorna (2012), "Student Assessment in Online Learning: Challenges and Effective Practices", **Journal of Online Learning and Teaching**, 8(3), 198-207.

Keh, Claudia (1990), "Feedback in the Writing Process: A Model and Methods for Implementation", **ELT Journal**, 44(4), 294-304.

Keleş, Pınar Ural and Aydın, Suleyman Aydın (2021), "University Students' Perceptions about Artificial Intelligence", **Shanlax International Journal of Education**, 9, 212-220.

Kellogg, Ronald et al. (2010), “Does Automated Feedback Help Students Learn to Write?”, **Journal of Educational Computing Research**, 42(2), 173-196.

Khoshnevisan, Babak and Le, Nhu (2018), “Augmented Reality in Language Education: A Systematic Literature Review”, **Adv. Glob. Educ. Res**, 2, 57-71.

Kim, Vivtoria (2018), “Technology-Enhanced Feedback on Student Writing in the English-Medium Instruction Classroom”, **English Teaching**, 73(4), 29-53.

Kirmizi, Özkan and Kiirmızı, Gülin Dağdeviren (2015), “An Investigation of L2 Learners' Writing Self-Efficacy, Writing Anxiety and Its Causes at Higher Education in Turkey”, **International Journal of Higher Education**, 4(2), 57-66.

Kiuhara, Sharlene et al. (2009), “Teaching Writing to High School Students: A National Survey”, **Journal of Educational Psychology**, 101(1), 136-160.

Klebanov, Beata Beigman et al. (2017), “Reflective Writing about the Utility Value of Science as a Tool for Increasing STEM Motivation and Retention-Can AI Help Scale Up?”, **International Journal of Artificial Intelligence in Education**, 27(4), 791-818.

Klekovkina, Vera and Denié-Higney, Laurence (2022), “Machine translation: Friend or Foe in the Language Classroom?”, **L2 Journal**, 14(1), 105-135.

Köbis, Nils and Mossink, Luca (2021), “Artificial Intelligence Versus Maya Angelou: Experimental Evidence That People Cannot Differentiate AI-Generated from Human-Written Poetry”, **Computers in Human Behavior**, 114, 106553.

Kocoń, Jan et al. (2023), “ChatGPT: Jack of all trades, master of none”, **Information Fusion**, (99) 101861.

Koenka, Alison and Anderman, Eric (2019), “Personalized Feedback as a Strategy for Improving Motivation and Performance Among Middle School Students”, **Middle School Journal**, 50(5), 15-22.

Koltovskaia, Svetlana (2020), “Student Engagement with Automated Written Corrective Feedback (AWCF) Provided by Grammarly: A Multiple Case Study”, **Assessing Writing**, 44, 100450.

Koole, Marguerite (2010), “Mobile Learning in Distance Education: Utility or Futility.” **Journal of Distance Education**, 24, (2) 59-82

Korkmaz, Gökben (2015), **The Effect of Creative Writing Method on the 6th Graders’ Perception of Writing Self Efficacy, Attitudes Towards Writing and Academic Success of Writing Skill**. Master's Thesis, Gaziantep University, Gaziantep.

Kormos, Judit and Kiddle, Thom (2013), “The Role of Socio-Economic Factors in Motivation to Learn English as a Foreign Language: The Case of Chile”, **System**, 41(2), 399-412.

Kukulska-Hulme, Agnes (2012), "How Should the Higher Education Workforce Adapt to Advancements in Technology for Teaching and Learning?", **The Internet and Higher Education**, 15(4), 247-254.

Kumar, Devesh (2023, July 7), "UK Universities Launch Probe After 400 Students Found Cheating Through ChatGPT", **Mint**. https://www.livemint.com/ai/artificial-intelligence/uk-universities-launch-probe-after-400-students-found-cheating-through-chatgpt-11688736355853.html (12.10.2023).

Kumar, Vivekanandan and Boulanger, David (2021), "Automated Essay Scoring and the Deep Learning Black Box: How are Rubric Scores Determined?", **International Journal of Artificial Intelligence in Education**, 31, 538-584.

Laird, John (2002), "Research in Human-Level AI Using Computer Games", **Communications of the ACM**, 45(1), 32-35.

Lamb, Martin and Arisandy, Fuziah Eka (2020), "The Impact of Online Use of English on Motivation To Learn", **Computer Assisted Language Learning**, 33(1-2), 85-108.

Lamprianou Iasonas et al. (2020), "The Longitudinal Stability of Rating Characteristics in an EFL Examination: Methodological and Substantive Considerations", **Language Testing**, 38(2), 273-301.

Lai, Yi-hsiu (2010), "Which do Students Prefer to Evaluate Their Essays: Peers or Computer Program", **British Journal of Educational Technology**, 41(3), 432-454.

Larson, Barbara et al. (2017), "Tools for Teaching Virtual Teams: A Comparative Resource Review", **Management Teaching Review**, 2(4), 333-347.

Leakey, Jonathan (2011), **Evaluating Computer-Assisted Language Learning**. Peter Lang AG, Washington.

Lee, James and VanPatten, Bill (2003), **Making Communicative Language Teaching Happen**. McGraw-Hill, New York.

Lee, Yong-Won et a008), "Analytic Scoring of TOEFL® CBT Essays: Scores from Humans and E-Rater®", **ETS Research Report Series**, 2008(1), i-71.

Legris, Paul et al. (2003), "Why Do People se Information Technology? A Critical Review of the Technology Acceptance Model", **Information & Management**, 40(3), 191-204.

Li, Qngyao et al. (2023), "Adapting Large Language Models for Education: Foundational Capabilities, Potentials, and Challenges". https://arxiv.org/abs/2401.08664 (12.10.2023).

Li, Rui (2022), "Understanding Foreign Language Writing Anxiety and its Correlates", **Frontiers in Psychology**, 13, 1031514.

Li, Rui et al. (2019), “Examining EFL learners’ individual antecedents on the adoption of automated writing evaluation in China”, **Computer Assisted Language Learning**, 32(7), 784-804.

Li, Yuwei (2024), **Towards a Contextualised Interpretation of Chinese University Student Engagement with Teacher Written Feedback,** Unpublished Doctoral dissertation, University of Hong Kong, China.

Li, Zhi et al. (2014), “The Role of Automated Writing Evaluation Holistic Scores in the ESL Classroom”, **System**, 44, 66-78.

Liao, Hui-Chuan (2016), “Using Automated Writing Evaluation to Reduce Grammar Errors in Writing”, **ELT Journal,** 70(3), 308-319.

Link, Stephanie et al. (2022), “Impact of Automated Writing Evaluation on Teacher Feedback, Student Revision, and Writing Improvement”, **Computer Assisted Language Learning**, 35(4), 605-634.

Little, David (1991), **Learner autonomy: Definitions, issues and problems,** Authentik, Dublin.

Liu, Meihua and Huang, Wenhong (2011), “An Exploration of Foreign Language Anxiety and English Learning Motivation”, **Education Research International**, 2011(1), 493167.

Liu, Ming et al. (2016), “Automated Essay Feedback Generation and its Impact on Revision”, **IEEE Transactions on Learning Technologies**, 10(4), 502-513.

Liu, Pengfei et al. (2023), “Pre-Train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing”, **ACM Computing Surveys**, 55(9), 1-35.

Liu, Yang and Lei, Ren (2022), “The Influence of Artificial Intelligence Technology on Teaching Under the Threshold of “Internet+”: Based on the Application Example of an English Education Platform”, **Wireless Communications and Mobile Computing**, 2022, 1-9.

Lo, Chung Kwan (2023), “What Is the Impact of Chatgpt on Education? A Rapid Review of the Literature”, **Education Sciences**, 13(4), 410.

Lo, Leo (2023), “The CLEAR Path: A Framework for Enhancing Information Literacy Through Prompt Engineering”, **The Journal of Academic Librarianship**, 49(4), 102720.

Long, Michael Hugh (1980), **Input, Interaction, and Second Language Acquisition,** Unpublished PhD Dissertation. University of California, Los Angeles.

Lou, Nigel Mantou and Noels, Kimberly (2018), “Western and Heritage Cultural Internalizations Predict EFL Students’ Language Motivation and Confidence”, **International Journal of Bilingual Education and Bilingualism**. 24(5), 636-650.

Luan, Hui et al. (2020), “Challenges and Future Directions of Big Data and Artificial Intelligence in Education”, **Frontiers in psychology**, 11, 580820.

Lyons, Zaza et al. (2020), "COVID-19 And the Mental Well-Being of Australian Medical Students: Impact, Concerns and Coping Strategies Used", **Australasian Psychiatry: Bulletin of Royal Australian and New Zealand College of Psychiatrists**, 28(6), 649–652.

Ma, Wei et al. (2023), **LLMs: Understanding Code Syntax and Semantics for Code Analysis** arXiv:2305.12138. (10.12.2023).

Mac Callum, Kathryn and Jeffrey, Lynn (2014), "Comparing the Role of ICT Literacy and Anxiety in the Adoption of Mobile Learning", **Computers in Human Behavior**, 39, 8-19.

MacIntyre, Peter (1999), "Language Anxiety: A Review of the Research for Language Teachers", Dolly Jesusita Young (Ed.), In **Affect in Foreign Language and Second Language Learning: A Practical Guide to Creating a Low Anxiety Classroom Atmosphere**, 24-45, McGraw-Hill Education, New York.

MacIntyre, Peter D., & Gardner, Robert C (1988), **The Measurement of Anxiety and Applications to Second Language Learning: An Annotated Bibliography**, Research Bulletin No. 672. University of Western Ontario, Department of Psychology. Sponsored by Social Sciences and Humanities Research Council of Canada.

MacIntyre, Peter and Gardner, Robert (1991), "Language Anxiety: Its Relationship to Other Anxieties and to Processing in Native and Second Languages", **Language Learnin**g, 41(4), 513-534.

Maes, Pattie (1995), "Artificial Life Meets Entertainment: Lifelike Autonomous Agents", **Communications of the ACM**, 38(11), 108-114.

Maghsudi, Setareh et al. (2021), "Personalized Education in The Artificial Intelligence Era: What to Expect Next", **IEEE Signal Processing Magazine**, 38(3), 37-50.

Mahadi, Tengku Sepora Tengku and Jafari, Sepideh Moghaddas Jafari (2012), "Motivation, Its Types, and Its Impacts in Language Learning", **International Journal of Business and Social Science**, 3(24), 230-235.

Manca, Stefania and Ranieri, Maria (2016), "Is Facebook Still a Suitable Technology-Enhanced Learning Environment? An Updated Critical Review of the Literature from 2012 To 2015", **Journal of Computer Assisted Learning**, 32(6), 503-528.

Marcus, Gary (2022, March 10), "Deep Learning is Hitting a Wall.", https://nautil.us/deep-learning-is-hitting-a-wall-238440/ (10.12.2023).

Marinescu, Valentina and Balica, Ecaterina (2013), "Korean Cultural Products in Eastern Europe: A Case Study of the K-Pop Impact in Romania", **Region: Regional Studies of Russia, Eastern Europe, and Central Asia**, 2(1), 113-135.

Marr, Bernard (2023, March 22), “ChatGPT: Whatare Hallucinations and why are They a Problem For AI Systems”, https://bernardmarr.com/chatgpt-what-are-hallucinations-and-why-are-they-a-problem-for-ai-systems/ (10.12.2023).

Matsumura, Shoichi and Hann, George (2004), “Computer Anxiety and Students' Preferred Feedback Methods in EFL Writing”, **The Modern Language Journal**, 88(3), 403-415.

Matthew, Kathryn et al. (2009), “Wiki as a Collaborative Learning Tool in a Language Arts Methods Class”, **Journal of Research on Technology in Education**, 42(1), 51-72.

McCarthy, John et al. (2006), “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence”, **AI magazine**, 27(4), 12-12.

McCroskey, James (1970), “Measures of Communication-Bound Anxiety”, **Speech Monographs**, 37, 269–277.

McGarrell, Hedy and Verbeem, Jeff (2007), “Motivating Revision of Drafts Through Formative Feedback”, **ELT Journal**, 61(3), 228-236.

McKenna, Nick et al. (2023), “Sources of Hallucination by Large Language Models on Inference Tasks”, https://arxiv.org/pdf/2305.14552 (11.12.2023)

McMurtry, Kim (2001), “E-cheating: Combating a 21st Century challenge”, **T. H. E. Journal**, 29(4), 36-38

Mede, Enisa and Karaırmak, Özlem (2017), “The Predictor Roles of Speaking Anxiety and English Self Efficacy on Foreign Language Speaking Anxiety”, **Journal of Teacher Education and Educators**, 6(1), 117-131.

Megahed, Fadel et al. (2023), “How Generative AI Models Such as Chatgpt Can e (Mis) Used in Spc Practice, Education, and Research? An Exploratory Study”, **Quality Engineering**, 36(2) 287-315.

vMeishar-Tal, Hagit (2015), “Teachers' Use of Wikipedia with Their Students”, **Australian Journal of Teacher Education (Online)**, 40(12), 126-140.

Mendi, Hatice Bahar (2009), **The Relationship Between Reading Strategies, Motivation and Reading Test Performance in Foreign Language Learning**, Unpublished Doctoral Dissertation, Marmara University, Turkey.

Microsoft (n.d.), “Microsoft Editor Checks Grammar And More in Documents, Mail, and the Web”, **Microsoft Support**. https://support.microsoft.com/en-us/office/microsoft-editor-checks-grammar-and-more-in-documents-mail-and-the-web-91ecbe1b-d021-4e9e-a82e-abc4cd7163d7 (10.12.2023).

Miles, Matthew et al. (1994), **Qualitative data analysis: An expanded sourcebook**, **Sage**, New York.

CMiller, Douglas and Brown, Eric (2018), "Artificial Intelligence in Medical Practice: The Question to the Answer?", **The American Journal of Medicine**, 131(2), 129-133.

Min, Lau Sing and Rahmat, Nurhazlini (2014), "English Language Writing Anxiety Among Final Year Engineering Undergraduates in University Putra Malaysia", **Advances in Language and Literary Studies**, 5(4), 102-106.

Mittelstadt, Brent Daniel et al. (2016), "The Ethics of Algorithms: Mapping the Debate", **Big Data & Society**, 3(2), 2053951716679679.

Mohsen, Mohammed Ali (2022), "Computer-Mediated Corrective Feedback to Improve L2 Writing Skills: A Meta-Analysis", **Journal of Educational Computing Research**, 60(5), 1253-1276.

Mohsen, Mohammed Ali and Abdulaziz, Alshahrani (2019), "The Effectiveness af Using a Hybrid Mode of Automated Writing Evaluation System on EFL Students' Writing", **Teaching English With Technology**, 19(1), 118-131.

Morphy, Paul and Graham, Steve (2012), "Word Processing Programs and Weaker Writers/Readers: A Meta-Analysis of Research Findings", **Reading and Writing**, 25(3), 641-678.

Murray Nick et al. (2020), "Artificial Intelligence to Diagnose Ischemic Stroke and Identify Large Vessel Occlusions: A Systematic Review", **Journal of Neurointerventional Surgery**, 12(2), 156-164.

Murthy, Sahana et al. (2017), "Technology Integration for Student-Centered Learning: A Model for Teacher Professional Development Programs", Siu Cheung Kong et al. (Eds.), In **Emerging Practices in Scholarship of Learning and Teaching in a Digital Era**, (55-74), Springer, Germany.

Nacak, Ayşegül et al. (2020), "Teacher Candidate Views on the Use of Youtube for Educational Purposes", **Online Journal of Communication and Media Technologies**, 10(2), e202003.

Nazari, Nabi et al. (2021), "Application of Artifcial Intelligence Powered Digital Writing Assistant in Higher Education: Randomized Controlled Trial", **Heliyon**, 7(5), e07014.

Negari, Giti Mousapour (2011), A Study on Strategy Instruction and EFL Learners' Writing Skill", **International Journal of English Linguistics**, 1(2), 299-307.

Nelson, Gayle and Carson, Joan (1998), "ESL Students' Perceptions of Effectiveness in Peer Response Groups", **Journal of Second Language Writing**, 7(2), 113-131.

Nelson, Gayle and Murphy, John (1993), "Peer Response Groups: Do L2 Writers Use Peer Comments in Revising Their Drafts?", **TESOL Quarterly**, 27(1), 135-141.

Ngo, Thu Huong (2015), **An Investigation into Students' Motivation to Learn English in Higher Education in Vietnam**, Unpublished Doctoral dissertation, Queensland University of Technology, Brisbane.

Ngo, Thuy Thi-Nhu et al. (2022), "The Effectiveness of Automated Writing Evaluation in EFL/ESL Writing: A Three-Level Meta-Analysis", **Interactive Learning Environments**, 32(2), 727-744.

Nguyen, Hong Thu Thi (2023), "Unproctored Assignment-Based Online Assessment in Higher Education: Stakeholder Evaluation of Issues", **Issues in Educational Research**, 33(1), 207-226.

Noels, Kimberly et al. (1999), "Perceptions of Teachers' Communicative Style and Students' Intrinsic and Extrinsic Motivation", **The Modern Language Journal**, 83(1), 23-34.

Noeth, Richard and Kobrin, Jennifer (2007), "Writing Changes in the Nation's K-12 Education System", **Research Notes**-34. Iowa City, IA: The College Board.

Nunes, Aandreia et al. (2022), "Effectiveness of Automated Writing Evaluation Systems in School Settings: A Systematic Review of Studies From 2000 To 2020", **Journal of Computer Assisted Learning**, 38(2), 599-620.

O'Connor, Siobhan and ChatGPT (2023), "Open Artificial Intelligence Platforms in Nursing Education: Tools for Academic Progress or Abuse?", **Nurse Education in Practice**, 66, 103537.

Öğüt, Banu (2022), **The Impact of Automated Corrective Feedback Tools on EFL Learners' Writing Performance**. Master's dissertation, Bahçeşehir University, İstanbul.

Özgür, Burcu and Griffiths, Carol (2013), "Second Language Motivation", **Procedia-Social and Behavioral Sciences**, 70, 1109-1114.

OpenAI (2021), "OpenAI GPT-3", https://openai.com/blog/gpt-3-apps/ (10.12.2023)

OpenAI (2023, March 23), "ChatGPT plugins", Retrieved from https://openai.com/blog/chatgpt-plugins#OpenAI (10.12.2023)

OpenAI (2023, April 5), "Our Approach to AI Safety", https://openai.com/blog/our-approach-to-ai-safety (10.12.2023)

Oral, Belgin et al. (2024), "University Students' States of Having Depressive Symptoms and Hopefulness About Getting Employed After Graduation", **Maltepe Tıp Dergisi**, 16(1), 7-14.

Oxford, Rebecca and Shearin, Jill (1994), "Language Learning Motivation: Expanding the Theoretical Framework", **The Modern Language Journal**, 78(1), 12-28.

Özer, Ömer and Akçayoğlu, Duygu İşpınar (2021), "Examining The Roles Of Self-Efficacy Beliefs, Self-Regulated Learning and Foreign Language Anxiety in the Academic Achievement of Tertiary EFL Learners", **Participatory Educational Research**, 8(2), 357-372.

Page, Ellis (1966), "The Imminence of... Grading Essays by Computer", **The Phi Delta Kappan**, 47(5), 238-243.

_____ (1994), “Computer Grading of Student Prose, Using Modern Concepts and Software”, **The Journal of Experimental Education**, 62(2), 127-142.

Palermo, Corey and Thomson, Margareta Maria (2018), “Teacher Implementation of Self-Regulated Strategy Development with an Automated Writing Evaluation System: Effects on the Argumentative Writing Performance of Middle School Students”, **Contemporary Educational Psychology**, 54, 255-270.

Pallant, Jullie (2020), **SPSS Survival Manual: A Step by Step Guide to Data Analysis Using IBM SPSS**, Routledge, London.

Pandita, Alka and Kiran, Ravi (2023), “The Technology Interface and Student Engagement are Significant Stimuli in Sustainable Student Satisfaction”, **Sustainability**, 15(10), 7923.

Patton, Michael Quinn (2002), “Two Decades of Developments in Qualitative Inquiry: A Personal, Experiential Perspective”, **Qualitative Social Work**, 1(3), 261-283.

Perelman, Les (2014), “When “The State of the Art” is Counting Words”, **Assessing Writing**, 21, 104-111.

Peritz, Aki (2022), “A.I. Is Making it Easier than Ever for Students to Cheat. SLATE”, **Future Tense**. The Slate Group. https://slate.com/technology/2022/09/ai-students-writing-cheating-sudowrite.html (10.12.2023).

Pero, James (2023, April 12), “Meet the Jailbreakers Hypnotizing ChatGPT into Bomb-Building”, https://www.inverse.com/tech/chatgpt-jailbreakers-reddit-open-ai-chatbot (10.12.2023).

Perrone, Lisa and Vickers, Margaret (2003), “Life After Graduation as a “Very Uncomfortable World”: An Australian Case Study”, **Education+ Training**, 45(2), 69-78.

Peterson, Mark (2016), **Computer Games and Language Learning**, Springer, New York.

Phuket, Pimpise Rattanailok Na and Othman, Normah Binti (2015), “Understanding EFL Students' Errors in Writing”, **Journal of Education and Practice**, 6(32), 99-106.

Plass, Jan et al. (2015), “Foundations of Game-Based Learning”, **Educational Psychologist**, 50(4), 258-283.

Polly, Drew and Byker, Erik (2020), “Considering the Role of Zone of Proximal Development and Constructivism in Supporting Teachers’ TPACK and Effective use of Technology”, **Revista de Educación a Distancia (RED)**, 20(64).

Powers, Donald et al. (2002), “Stumping E-Rater: Challenging the Validity of Automated Essay Scoring”, **Computers in Human Behavior**, 18(2), 103-134.

Ramesh, Dadi and Sanampudi, Suresh Kumar (2022), “An Automated Essay Scoring Systems: A Systematic Literature Review”, **Artificial Intelligence Review**, 55(3), 2495-2527.

Ranalli, Jim and Yamashita, Taichi (2022), "Automated Written Corrective Feedback: Error-Correction Performance and Timing of Delivery", **Language Learning & Technology**, 26(1)1-25

Ratanasiripong, Paul et al. (2022), "Impact of Psychosocial Factors on Academic Performance of Nursing Students in Thailand", **Journal of Health Research,** 36(4), 598-606.

Riasati, Mohammad Javad (2011), "Language Learning Anxiety from EFL Learners' Perspective", **Middle-East Journal of Scientific Research**, 7(6), 907-914.

Rosen, Larry (2013), "Facebook and Texting Made Me Do it: Media-Induced Task-Switching While Studying", **Computers in Human Behavior**, 29(3), 948-958.

Rudolph, Jürgen et al. (2023), "ChatGPT: Bullshit Spewer or the End of Traditional Assessments in Higher Education?", **Journal of Applied Learning and Teaching**, 6(1) 342-263.

Russell, Anne (1995), "Stages in Learning New Technology: Naive Adult Email Users", **Computers & Education**, 25(4), 173-178.

Rückert, Hans-Werner (2015), "Students′ Mental Health and Psychological Counselling in Europe", **Mental Health & Prevention**, 3(1-2), 34-40.

Sağlamel, Hasan, (2018), **A Comparison of the Impact of Teacher Feedback Within and Irrespective of the English as a Foreign Language Learners' Zone of Proximal Development**, Unpublished doctoral dissertation, Karadeniz Technical University, Institute of Social Sciences, Trabzon.

Sağlamel, Hasan, and Kayaoğlu, Mustafa Naci (2015), "English Major Students' Perceptions of Academic Writing: A Struggle Between Writing to Learn and Learning to Write", **Journal of History Culture and Art Research**, 4(3) 37-52.

Sanders, Micheal (2005), **The Effect of Immediate Feedback and After Action Reviews (AARS) on Learning, Retention and Transfer**, http://purl.fcla.edu/fcla/etd/CFE0000441 (10.12.2023).

Sarason, Irwin and Ganzer, Victor (1962), "Anxiety, Reinforcement, and Experimental Instructions in a Free Verbalization Situation", **Journal of Abnormal and Social Psychology**, 65, 300–307.

Sarman, Gülpınar (2022), **EFL teachers' Perceptions on Task-Based Language Teaching in the Context of a Preparatory School**, Unpublished Master's thesis, Bursa Uludağ University

Savignon, Sandra (1991), "Communicative Language Teaching: State of the Art", **TESOL Quarterly**, 25(2), 261-278.

Schwarz, Barıuch and Galit, Caduri (2016), "Novelties in the Use of Social Networks by Leading Teachers in Their Classes", **Computers & Education**, 102, 35-51.

Selwyn, Neil (2016), "Digital downsides: Exploring University Students' Negative Engagements with Digital Technology", **Teaching in Higher Education**, 21(8), 1006-1021.

Sevinç, Yeşim (2018), "Language Anxiety in the Immigrant Context: Sweaty Palms?", **International Journal of Bilingualism**, 22(6), 717-739.

Shaaban, Kassim and Ghaith, Ghazi (2000), "Student Motivation to Learn English as a Foreign Language", **Foreign Language Annals**, 33(6), 632-644.

Shang, Hui-Fang (2022), "Exploring Online Peer Feedback and Automated Corrective Feedback on EFL Writing Performance", **Interactive Learning Environments**, 30(1), 4-16.

Sheen, Younghee (2007), "The Effect of Focused Written Corrective Feedback and Language Aptitude on ESL Learners' Acquisition of Articles", **TESOL Quarterly**, 41(2), 255-283.

Shell, Duane et al. (1989), "Self-Efficacy and Outcome Expectancy Mechanisms Inreading and Writing Achievement", **Journal of Educational Psychology**, 81, 91–100.

Shermis, Mark (2014), "State-Of-The-Art Automated Essay Scoring: Competition, Results, and Future Directions From a United States Demonstration", **Assessing Writing**, 20, 53-76.

Shermis, Mark and Burstein, Jill (2003), **Automated Essay Scoring: A Cross-Disciplinary Perspective**, Routledge, New York.

Shermis, M. et al. (2016), "Automated Writing Evaluation: A Growing Body of Knowledge", Charles MacArthur, Steve Graham and Jill Fitzgerald (Eds.), In **Handbook of Writing Research (**395–409), Guilford Press, New York.

Shermis, Mark and Hamner, Ben (2013), "Contrasting State-Of-The-Art Automated Scoring of Essays". Mark D. Shermis and Jill Burstein, In **Handbook of Automated Essay Evaluation (**313-346), Routledge.

Shermis, Mark et al. (2002), "Trait Ratings for Automated Essay Grading", **Educational and Psychological Measurement**, 62(1), 5-18.

Shimbun, Yomiuri (2023, April 26), "Chatgpt can be Tricked into Generating Malware, Bomb-Making Instructions", https://asianews.network/chatgpt-can-be-tricked-into-generating-malware-bomb-making-instructions/ (10.12.2023).

Shute, Valerie (2008), "Focus on formative feedback", **Review of Educational Research**, 78(1), 153-189.

Skar, Gustaf Bernhard Uno et al. (2021), "Learning Loss During The COVID-19 Pandemic and the Impact of Emergency Remote Instruction on First Grade Students' Writing: A Natural Experiment", **Journal of Educational Psychology**, 114(7) 1553-1566.

Skehan, Peter (1991), "Individual Differences in Second Language Learning", **Studies in Second Language Acquisition,** 13(2), 275-298.

Smutny, Pavel and Schreiberova, Petra (2020), "Chatbots for Learning: A Review of Educational Chatbots for the Facebook Messenger", **Computers & Education**, 151, 103862.

Stevenson, Marie and Phakiti, Aek (2014), "The Effects of Computer-Generated Feedback on the Quality of Writing", **Assessing Writing**, 19, 51–65.

Stevenson, Marie and Phakiti, Aek (2019), "Automated Feedback and Second Language Writing", Ken Hyland & Fiona Hyland (Eds.), In **Feedback in Second Language Writing** (pp. 125–142), Cambridge University Press, Cambridge.

Stockwell, Glenn (2010), "Using Mobile Phones for Vocabulary Activities: Examining the Effect of the Platform", **Language Learning & Technology**, 16(3), 23-39.

_____ (2012), **Computer-Assisted Language Learning: Diversity in Research and Practice**, Cambridge University Press, Cambridge.

Stokel-Walker, Chris (2023, January 18), "ChatGPT Listed as Author on Research Papers: Many Scientists Disapprove", **Nature,** 613(7945), 620-621

Stone, Philip et al. (1966), **The General Inquirer: A Computer Approach to Content Analysis**, **M. I. T. Press, Massachusetts.**

Strzelecki, Artur (2023), "To Use or Not to Use ChatGPT in Higher Education? A Study of Students' Acceptance and Use of Technology", **Interactive Learning Environments**, 31, 1-14.

Subaşı, Gonca (2010), "What Are the Main Sources of Turkish EFL Students' Anxiety in Oral Practice?", **Turkish Online Journal of Qualitative Inquiry**, 1(2), 29-49.

Sucan, Serdar (2019), "The Relationship Between Hope and Perceived Stress in Teacher Candidates", **International Journal of Higher Education**, 8(2), 1-6.

Surahman, Ence and Wang, Tzu-Hua (2022), "Academic Dishonesty and Trustworthy Assessment in Online Learning: A Systematic Literature Review", **Journal of Computer Assisted Learning**, 38(6), 1535-1553.

Syed Abdul, Shabbir et al. (2022), "Virtual Reality Enhancing Medical Education and Practice: Brief Communication", **Digital Health**, 8, 20552076221143948.

Tai, Tzu-Yu and Chen, Howard Hao-Jan (2023), "The Impact of Google Assistant on Adolescent EFL Learners' Willingness to Communicate", **Interactive Learning Environments**, 31(3), 1485-1502.

Talan, Tarık and Kalınkara, Yusuf (2023), "Yükseköğrenimde Yapay Zekânın Rolü: Anatomi Dersi için ChatGPT Değerlendirmesi", **Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi**, 7(1), 33-40.

Tasdemir, Muhammed Salih and Yalcin Arslan, Fadime (2018), "Feedback Preferences of EFL Learners With Respect to Their Learning Styles", **Cogent Education**, 5(1), 1481560.

Temsah, Omar et al. (2023), "Overview of Early Chatgpt's Presence in Medical Literature: Insights from a Hybrid Literature Review by Chatgpt and Human Experts", **Cureus**, 15(4) 1-7.

Tay, Yi et al. (2022), "Efficient transformers: A survey", **ACM Computing Surveys,** 55(6), 1–28.

Thi, Nang. Kham et al. (2023), "Higher-Proficiency Students' Engagement with and Uptake of Teacher and Grammarly Feedback in an EFL Writing Course", **Innovation in Language Learning and Teaching**, 17(3), 690-705.

Tobar-Muñoz, Hendrys et al. (2017), "Augmented Reality Game-Based Learning: Enriching Students' Experience During Reading Comprehension Activities", **Journal of Educational Computing Research,** 55(7), 901-936.

Topsakal, Oguzhan and Topsakal, Elif (2022), "Framework for a Foreign Language Teaching Software for Children Utilizing AR, Voicebots and ChatGPT (Large Language Models)", **The Journal of Cognitive Systems**, 7(2), 33-38.

Tran, Alisia et al. (2018), "Financial Stress, Social Supports, Gender, and Anxiety During College: A Stress-Buffering Perspective", **The Counseling Psychologist**, 46(7), 846-869.

Tsai, Ya-Chin and Li, Yi-Chih (2012), "Test Anxiety and Foreign Language Reading Anxiety in a Reading-Proficiency Test", **Journal of Social Sciences**, 8(1), 95-103

Tunçel, Hayrettin (2015), "The Relationship Between Self-Confidence and Learning Turkish as a Foreign Language", **Educational Research and Reviews**, 10(18), 2575-2589.

Turhan, Nilhan Sölpük (2020), "Gender Differences in Academic Motivation: A Meta-Analysis", **International Journal of Psychology and Educational Studies**, 7(2), 211-224.

Türnük, Tülin and Aydın, Selami (2020), "Interlanguage Effect on Writing Anxiety in a Foreign Language Context", **The Literacy Trek**, 6(1), 1-19.

Upadhyay, Ashwani Kumar and Khandelwal, Komal (2019), "Artificial Intelligence-Based Training Learning from Application", **Development and Learning in Organizations: An International Journal**, 33(2), 20-23.

Uzun, Kutay and Köksal, Handan (2020), "Direct vs Indirect Written Corrective Feedback: An Action Research", **Trakya Eğitim Dergisi**, 10(1), 169-182.

Uzun, Levent (2023), "Chatgpt and Academic Integrity Concerns: Detecting Artificial Intelligence Generated Content", **Language Education and Technology**, 3(1), 45-54.

Vallerand, Robert. et al. (1992), "The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education", **Educational and Psychological Measurement**, 52(4), 1003-1017.

Van Dis, Eva et al. (2023), "ChatGPT: five priorities for research", **Nature**, 614(7947), 224-226.

Vashishth, Tarun Kumar et al. (2024), "AI-driven learning analytics for personalized feedback and assessment in higher education". Tien V. T. Ngguyen and Nhut T. M. Vo (Eds.), In **Using Traditional Design Methods to Enhance AI-Driven Decision Making** (206-230), IGI Global, Pennsylvania.

Vaswani, Aashish et al. (2017), "Attention is All You Need", **Advances in Neural Information Processing Systems,** 30(1), 261-272.

Venkatesh, Viswanath et al. (2012), "Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance And use of Technology", **MIS Quarterly**, 157-178.

Vojak, Colleen et al. (2011), "New Spaces and Old Places: An Analysis of Writing Assessment Software", **Computers and Composition**, 28(2), 97-111.

Vygotsky, Lev Semyonovich & Cole, Michael (1978), **Mind in Society: Development of Higher Psychological Processes**, Harvard University Press, Harvard.

Wang, Charles Xiaoxue et al. (2012), "Learning Effects of an Experimental EFL Program in Second Life", **Educational Technology Research and Development**, 60, 943-961.

Wang, Yuntao et al. (2023), "A Survey on Chatgpt: AI-Generated Contents, Challenges, and Solutions", **IEEE Open Journal of the Computer Society** 4, 280-302.

Wang, Ying-Jian et al. (2013), "Exploring the Impact of Using Automated Writing Evaluation in English as a Foreign Language University Students' Writing", **Computer Assisted Language Learning**, 26(3), 234–257.

Ware, Paige (2014), "Feedback for Adolescent Writers in the English Classroom", **Writing & Pedagogy**, 6(2), 223-249.

Warschauer, Mark (1996), "Computer-Assisted Language Learning: An Introduction", Sandra Fotos (Ed.), In **Multimedia Language Teaching**. Tokyo: Los International.

Warschauer, Mark (2002), "A Developmental Perspective on Technology in Language Education", **TESOL Quarterly**, 36(3), 453-475.

Warschauer, Mark and Grimes, Douglas (2008), "Automated Writing Assessment in the Classroom" **Pedagogies: An International Journal**, 3(1), 22-36.

Warschauer, Mark and Healey, Deborah (1998), "Computers and Language Learning: An Overview", **Language Teaching,** 31(2), 57-71.

Warschauer, Mark and Ware, Paige (2006), "Automated Writing Evaluation: Defining the Classroom Research Agenda", **Language Teaching Research**, 10(2), 157-180.

Weber-Wulff, D., et al. (2023), "Testing of detection tools for AI-generated text", **International Journal for Educational Integrity**, 19(1), 26.

Wei, M. (2015), "An Empirical Study into Effects of Feedback on College English Writing", **Journal of Tianjin Foreign Studies University**, 22, 43-50.

Weigle, Sara C. (2013), "English as a Second Language Writing and Automated Essay Evaluation", Mark D. Shermis & Jill Burstein (Eds.), In **Handbook of Automated Essay Evaluation** (58-76), Routledge, New York.

Welbers, Kasper et al. (2019), "Gamification as a Tool for Engaging Student Learning: A Field Experiment with a Gamified App", **E-learning and Digital Media**, 16(2), 92-109.

White, Robert. W. (1959), "Motivation Reconsidered: The Concept of Competence", **Psychological Review**, 66(5), 297-333.

Wiggers, Kyle (2022, December 17), "OpenAI's Attempts to Watermark AI Text Hit Limits", https://techcrunch.com/2022/12/10/openais-attempts-to-watermark-ai-text-hit-limits/?guccounter=1 (10.12.2023).

Williamson, David et al. (2012), "A Framework for Evaluation and use of Automated Scoring", **Educational Measurement: Issues and Practice**, 31(1), 2–13.

Wilson, Amanda and Lengeling, M. Martha (2021), "Language Learning in the Time of COVID-19: ELT Students' Narrated Experiences in Guided Reflective Journals", **Íkala**, 26(3), 571-585.

Wilson, Joshua (2017), "Associated Effects of Automated Essay Evaluation Software on Growth in Writing Quality for Students with and Without Disabilities", **Reading and Writing**, 30, 691-718.

Wilson, Joshua and Czik, Amanda (2016), "Automated Essay Evaluation Software in English Language Arts Classrooms: Effects on Teacher Feedback, Student Motivation, and Writing Quality", **Computers & Education**, 100, 94-109.

Wilson, Joshua et al. (2021), "Automated Feedback and Automated Scoring in the Elementary Grades: Usage, Attitudes, and Associations with Writing Outcomes in a Districtwide Implementation of MI Write", **International Journal of Artificial Intelligence in Education**, 31(2), 234-276.

Wilson, Joshua and Rodrigues, Jessica (2020), "Classification Accuracy and Efficiency of Writing Screening Using Automated Essay Scoring", **Journal of School Psychology**, 82, 123-140.

Wilson, Joshua and Roscoe, Rod D. (2020), "Automated Writing Evaluation and Feedback: Multiple Metrics of Efficacy", **Journal of Educational Computing Research**, 58(1), 87-125.

Wong, Ruth M. (2014), "Motivation to Learn English and School Grade Level: The Case of Newly Arrived Hong Kong Students", **Porta Linguarum**, 21, 37–50, 23821205211024078.

Xi, Xiaoming (2010), "Automated Scoring and Feedback Systems: Where are We and Where are We Heading?", **Language Testing**, 27(3), 291-300.

Xiao, Yangyu and Zhi, Yuying (2023), “An Exploratory Study of EFL Learners’ Use of Chatgpt for Language Learning Tasks: Experience and Perceptions”, **Languages**, 8(3), 212.

Xu, Yongjun., et al. (2021), “Artificial Intelligence: A Powerful Paradigm for Scientific Research”, **The Innovation**, 2(4), 1-21.

Yaghi, Hussein et al. (1998), “Teachers' Computer Anxiety: An International Perspective”, **Computers in Human Behavior**, 14(2), 321-336.

Yalçın, Hüsna (2018), **Üniversite Hazırlık Sınıfı Öğrencilerinin Ders Dışı Etkinliklere Katılımları ile Yabancı Dil Öğrenmeye Güdülenmeleri Arasındaki Ilişki**, Unpublished Master's thesis, Ankara Üniversitesi, Ankara.

Yan, Da (2023), “Impact of Chatgpt on Learners in a L2 Writing Practicum: An Exploratory Investigation”, **Education and Information Technologies**, 28(11), 13943-13967.

Yang, Lily Wei Yun et al. (2021), “Deep Learning-Based Natural Language Processing in Ophthalmology: Applications, Challenges and Future Directions”, **Current Opinion in Ophthalmology**, 32(5), 397-405.

Yang, Yu-Fen (2010), “Students’ Reflection on Online Self-Correction and Peer Review to Improve Writing”, **Computers & Education**, 55(3), 1202-1210.

Yashima, Tomoko (2002), “Willingness to communicate in a second language: The Japanese EFL context”, **The Modern Language Journal**, 86(1), 54-66.

Yastıbaş, Gülşah Çınar and Yastıbaş, Ahmet Erdost (2015), “The Effect of Peer Feedback n Writing Anxiety In Turkish EFL Students”, **Procedia-Social and Behavioral Sciences**, 199, 530-538.

Yeh, Yu-li et al. (2007), “The Influence of Automatic Essay Evaluation and Bilingual Concordancing on EFL Students' Writing”, 英語教學期刊, 31(1), 117-160.

Yıkılmaz, Halil (2020), **Foreign Language Learning Motivation: A Study at Atılım University Preparatory School,** Unpublished Master's thesis, Ufuk University, Ankara.

Yoestara, Marisa and Putri, Zaiyana (2019), “PODCAST: An Alternative Way to Improve EFL Students’ Listening and Speaking Performance”, **Englisia**: **Journal of Language, Education, and Humanities**, 6(1), 15-26.

Yüncü Kurt, Pınar (2014), **The Effects of ARCS Motivational Model on Student Motivation to Learn English,** Unpublished Master's thesis, Anadolu University, Eskişehir.

Zerey, Özge Gül (2013), “Pre-Service EFL Teachers’ Foreign Language Writing Anxiety: Some Associated Factors”, **Dil Dergisi**, (160), 42-65.

Zhai, Na and Ma, Xiaomei (2022), “Automated writing evaluation (AWE) feedback: A systematic investigation of college students’ acceptance”, **Computer Assisted Language Learning**, 35(9), 2817-2842.

_____ (2023), "The Effectiveness of Automated Writing Evaluation on Writing Quality: A Meta-Analysis", **Journal of Educational Computing Research**, 61(4), 875-900.

Zhang, Zhe and Hyland, Ken (2023), "The Role of Digital Literacy in Student Engagement with Automated Writing Evaluation (AWE) Feedback on Second Language Writing", **Computer Assisted Language Learning**, 1-26.

Zhang, Zhongyu (2024), "New Communicative Language Teaching Methods: How ChatGPT is Used in English Teaching and Its Impacts", **Journal of Education, Humanities and Social Sciences**, 32, 74-78.

Zhou, Cong (2023), "Integration of Modern Technologies in Higher Education on the Example of Artificial İntelligence Use", **Education and Information Technologies**, 28(4), 3893-3910.

Zhou, Tongquan et al. (2023), "Chinese Intermediate English Learners Outdid ChatGPT in Deep Cohesion: Evidence from English Narrative Writing", **System**, 118, 103141.

Zhou, Yishu (2013), "A Comparative Study of Teacher Feedback, Peer Feedback and Online Feedback in English Writing", **Foreign Language World**, 3, 87-96.

Zuckerman, Marvin (1960), "The Development of an Affect Adjective Check List for the Measurement of Anxiety", **Journal of Consulting Psychology**, 24, 457–462.

APPENDICES

Appendix A: Student Questionnaire

Merhaba,

Bu anket, yüksek lisans tezi çalışması kapsamında yabancı dil bölümündeki İngilizce yazma dersi alan öğrencilere yönelik hazırlanmıştır. Çalışma, İngilizce yazma ödevlerine öğretmen geri bildirimi ve otomatik geri bildirimin öğrenci yazma kaygısı ve motivasyonuna etkilerini incelemeyi ve hafifletmeyi amaçlamaktadır. Lütfen aşağıdaki soruları kendinize uyacak şekilde doğru cevaplayınız. Verdiğiniz cevaplar, gizlilik esas alınarak analiz edilip çalışmada kullanılmak üzere kullanılacaktır. Ankette sağladığınız kişisel bilgilerin herhangi biri kamuya açık şekilde sunulmayacaktır. Herhangi bir noktada çalışmaya katılımdan vazgeçebilir, çalışma hakkında ekli eposta adresinden bilgi alabilirsiniz.

Katılımınız için teşekkür ederim.

****

E-posta: **********

Appendix A.1
Yabancı Dil Yazma Kaygı Ölçeği

Yaş:

Cinsiyet:

Sınıf:

	Aşağıdaki ifadeleri dikkatle okuyunuz. Her satırdaki ifadenin size ne kadar uygun olduğuna göre 1 "Hiç katılmıyorum", 5 de "Tamamen katılıyorum" olacak şekilde yuvarlak ile işaretleyin.	**Hiç Katılmıyorum**	**Katılmıyorum**	**Kararsızım**	**Katılıyorum**	**Tamamen Katılıyorum**
1	İngilizce yazarken endişelenirim.	**1**	**2**	**3**	**4**	**5**
2	İngilizce yazarken zaman sınırlaması olduğunda endişelenirim.	**1**	**2**	**3**	**4**	**5**
3	İngilizce yazarken değerlendirileceğimi bilirsem endişelenirim.	**1**	**2**	**3**	**4**	**5**
4	Fikirlerimi önce ana dilde yazar sonra İngilizceye çeviririm	**1**	**2**	**3**	**4**	**5**
5	Mecbur kalmadıkça İngilizce yazmam.	**1**	**2**	**3**	**4**	**5**
6	Sınavda İngilizce yazmaya başladığım zaman aklıma hiçbir şey gelmez.	**1**	**2**	**3**	**4**	**5**
7	İngilizce yazılarımın diğer arkadaşlarıma kıyasla daha kötü olmasından endişe ediyorum.	**1**	**2**	**3**	**4**	**5**
8	Yazılarım değerlendirildiğinde düşük bir not almaktan kaygılanırım.	**1**	**2**	**3**	**4**	**5**
9	Sosyal medya gibi ortamlarda İngilizce yazmaktan kaçınırım.	**1**	**2**	**3**	**4**	**5**
10	İngilizce yazarken zaman sınırlaması olursa düşüncelerimi toparlayamam.	**1**	**2**	**3**	**4**	**5**
11	Yazarken İngilizceyi tercih ederim.	**1**	**2**	**3**	**4**	**5**

Appendix A.1: (Continued)

12	Başkalarının İngilizce yazılarımı okurken alay edeceğinden korkarım.	**1**	**2**	**3**	**4**	**5**
13	Hazırlıksız İngilizce paragraf yazmam istenirse takılıp kalırım.	**1**	**2**	**3**	**4**	**5**
14	İngilizce yazmaktan kaçınırım.	**1**	**2**	**3**	**4**	**5**
15	Başkalarının yazılarım hakkında ne düşüneceğini umursamam.	**1**	**2**	**3**	**4**	**5**
16	Sınıfta tartışılmak üzere örnek olarak benim yazımın seçilmesi beni tedirgin eder.	**1**	**2**	**3**	**4**	**5**
17	Ders dışında İngilizce yazmak için fırsatları değerlendiririm.	**1**	**2**	**3**	**4**	**5**
18	İngilizce yazarken genellikle gerilirim.	**1**	**2**	**3**	**4**	**5**

Appendix A.2

L2 Writing Motivation Scale

Yaş:

Cinsiyet:

Sınıf:

	Aşağıdaki ifadeleri dikkatle okuyunuz. Her satırdaki ifadenin size ne kadar uygun olduğuna göre 1 "Hiç katılmıyorum" 5 de "Tamamen katılıyorum" olacak şekilde yuvarlak ile işaretleyin.	**Hiç Katılmıyorum**	**Katılmıyorum**	**Kararsızım**	**Katılıyorum**	**Tamamen Katılıyorum**
1	İngilizce yazı yazarken hissettiklerimi kağıda döktüğüm için İngilizce yazı yazma isteğim artıyor.	**1**	**2**	**3**	**4**	**5**
2	İngilizce yazı yazarak yazma yeteneğim geliştiği için İngilizce yazma isteğim artıyor.	**1**	**2**	**3**	**4**	**5**
3	Yeni şeyler öğrenmekten keyif aldığım için yazdıkça İngilizce yazma isteğim artıyor.	**1**	**2**	**3**	**4**	**5**
4	Kendime başarılı bir insan olduğumu kanıtlamak istediğim için İngilizce yazma isteğim artıyor.	**1**	**2**	**3**	**4**	**5**
5	İngilizce yazma isteğim, İngilizce yazmayı sevdiğim için artıyor.	**1**	**2**	**3**	**4**	**5**
6	İngilizce dersinde başarısız olduğumda İngilizce yazma isteğim artıyor.	**1**	**2**	**3**	**4**	**5**
7	Derslere katıldıkça İngilizce yazma isteğim artıyor.	**1**	**2**	**3**	**4**	**5**
8	Ailem İngilizce bilmemi istediği için İngilizce yazma isteğim artıyor.	**1**	**2**	**3**	**4**	**5**
9	Teknolojiden uzak olduğumda İngilizce yazma isteğim artıyor.	**1**	**2**	**3**	**4**	**5**
10	İyi bir iş bulmama faydalı olacağını düşündüğüm için İngilizce yazma isteğim artıyor.	**1**	**2**	**3**	**4**	**5**
11	Dürüst olmak gerekirse bilmiyorum. Aslında İngilizce yazarak zamanımı boşa harcadığımı düşünüyorum.	**1**	**2**	**3**	**4**	**5**
12	Hiçbir fikrim yok. İngilizce yazarken ne yaptığımı anlamıyorum.	**1**	**2**	**3**	**4**	**5**
13	İngilizce yazmak bana uzak kalıyor.	**1**	**2**	**3**	**4**	**5**
14	İngilizce yazmayı anlamlı bulmuyorum.	**1**	**2**	**3**	**4**	**5**

ÖĞRENCİLERİN OTOMATİK GERİBİLDİRİM ÜZERİNE TUTUMLARINA DAİR MÜLAKAT SORULARI

Group A (Teacher Feedback)

1. Size İngilizce yazı yazma ödevi verildiğinde genel olarak hissettiğiniz duygular neler? (Mutlu, Hevesli, Stresli, Endişeli, Bıkkın, İsteksiz vb.)
2. Yazma ödevlerinizi teslim ederken öğretmenin size verdiği geri bildirimler hakkında ne düşünüyorsunuz? (bu ödevler yazma yeteneğinizi nasıl etkiliyor?)
3. Sizce etkin geri bildirim nedir, nasıl olmalıdır? (Sizin için ideal feedback nasıldır mesela?)
4. Yazma ödevlerine geri bildirim almak size nasıl hissettirdi?

4.1. Ödevlerinizin sınıfta öğretmen tarafından okunması veya düzeltilmesi size genel olarak nasıl hissettirir? (Korku, Endişe, Gurur vb.)

5. Hiç otomatik geri bildirim araçlarından yararlandınız mı? Bu deneyimden bahseder misiniz? Öğretmenden geri bildirim almaya kıyasla otomatik geri bildirimin sizce ne tür farkları vardır?
6. Size göre geri bildirimin eksik olduğu konular var mı? Varsa neler? Bu eksiklerin nasıl giderilebileceğini düşünüyorsunuz?
7. Aldığınız geri bildirimleri yeterli buluyor musunuz?
8. Yazınızda çıkan hataların sayısı sizi şaşırttı mı? Gördüğünüz bu düzeltmeler size nasıl hissettirdi? (bu düzeltmelere karşılık olarak neler yaptınız?)
9. Geri bildirim almak sizdeki yazma isteğini etkiledi mi? Nasıl?
10. Geri bildirim almak İngilizce yazma kaygınızı etkiler mi? Nasıl?
11. Yazma kaygısı hisseder misiniz? Nasıl? Ne zaman?

11.1. Yazma kaygınız varsa, azalması için neler yapılabilir? Yazma kaygınızın azaltılmasına en çok fayda sağlayan etmen nedir? (Neden?)

12. Genel olarak kendinizi yazmaya istekli hisseder misiniz? Sizi yazmaya istekli kılan durumlar nelerdir?

13. Yazma motivasyonunuzun artırılması için neler yapılabilir?
14. Yazma motivasyonunuzu en çok artıran etmenler nelerdir?
15. Aldığınız geri bildirimde en çok yararlı olduğunu düşündüğünüz kısım veya kısımlar hangisi veya hangileri? (Gramer yazım yanlışları vb. düzeltmeler veya içerik anlamındaki öneri ve yorumlar).
16. Yazılarınızın altında sağlanan yorumlar size nasıl bir deneyim sağladı?
17. Öğretmenin sağladığı geri bildirimin etkinliğini artırmak için neler yapılabilir?

Group B (Automated Feedback)

1. Size İngilizce yazı yazma ödevi verildiğinde genel olarak hissettiğiniz duygular neler? (Mutlu, Hevesli, Stresli, Endişeli, Bıkkın, İsteksiz vb.)

2. Yazma ödevlerinizi teslim ederken öğretmenin size verdiği geri bildirimler hakkında ne düşünüyorsunuz? (bu ödevler yazma yeteneğinizi nasıl etkiliyor?)
3. Sizce etkin geri bildirim nedir, nasıl olmalıdır? (Sizin için ideal feedback nasıldır mesela?)
4. Yazma ödevlerine otomatik geri bildirim almak size nasıl hissettirdi?

4.1. Ödevlerinizin sınıfta öğretmen tarafından okunması veya düzeltilmesi size genel olarak nasıl hissettirir? (Korku, Endişe, Gurur vb.)

5. Hiç otomatik geri bildirim araçlarından (grammarly, paperpal, quillbot vb.) yararlandınız mı? Bu deneyimden bahseder misiniz? Öğretmenden geri bildirim almaya kıyasla otomatik geri bildirimin sizce ne tür farkları var?
6. Size göre otomatik geri bildirimin eksik olduğu konular var mı? Varsa neler? Bu eksiklerin nasıl giderilebileceğini düşünüyorsunuz?
7. Aldığınız geri bildirimleri yeterli buluyor musunuz?
8. Yazınızda çıkan hataların sayısı sizi şaşırttı mı? Gördüğünüz bu düzeltmeler size nasıl hissettirdi?(bu düzeltmelere karşılık olarak neler yaptınız?)
9. Geri bildirim almak sizdeki yazma isteğini etkiledi mi? Nasıl?
10. Geri bildirim almak İngilizce yazma kaygınızı etkiler mi? Nasıl?
11. Yazma kaygısı hisseder misiniz? Nasıl? Ne zaman?

11.1. Yazma kaygınız varsa, azalması için neler yapılabilir? Yazma kaygınızın azaltılmasına en çok fayda sağlayan etmen nedir? (Neden?)

12. Genel olarak kendinizi yazmaya istekli hisseder misiniz? Sizi yazmaya istekli kılan durumlar nelerdir?
13. Yazma motivasyonunuzun artırılması için neler yapılabilir?
14. Yazma motivasyonunuzu en çok artıran etmenler nelerdir? (Follow-up: Otomatik geri bildirim mi, yoksa öğretmen geri bildirimi mi? Neden?)
15. Aldığınız feedbackte en çok yararlı olduğunu düşündüğünüz kısım veya kısımlar hangisi veya hangileri? l(Gramer yazım yanlışları vb. düzeltmeler veya içerik anlamındaki öneri ve yorumlar). Bu aldığınız formatta feedback almaya devam etmek ister miydiniz? Neden?
16. Yazılarınızın altında sağlanan yorumlar size nasıl bir deneyim sağladı?
17. Otomatik geri bildirimin etkinliğini artırmak için neler yapılabilir?

Appendix-C. Transcription Sample of the Semi-Structured Interviews

Researcher: İlk olarak, İngilizce yazma ödevi sana verildiğinde genel olarak hissettiğin duygular neler, ne gibi hissediyorsun? Mesela mutlu hevesli stresli endişeli bıkkın isteksiz falan bunların hepsi olabilir. Genel olarak nasıl hissediyorsun?

Participant A6M: Ya bu ödev verildiğinde ilk başta böyle bir stres oluyor ister istemez hani "yapabilir miyim" "acaba yazar mıyım", "ne yazayım" falan düşüncesi bir kaygısı oluyor ama hani eğer konu bana uygunsa içerik hani benim vakıf olduğum bir şey ise pek sıkıntı olmuyor. Yani kalemi aldığım zaman hani oturduğum zaman sandalyeye geliyor direkt hani, akıyor. O yüzden çok sıkıntı olmuyor.

Researcher: Anladım. Konuyla biraz daha alakası var diyorsun?

Participant A6M: Evet, yani içerik bana uygunsa bana hitap ediyorsa çok rahat olabiliyor.

Researcher: Anladım. Yazma ödevlerini teslim ederken öğretmenin verdiği geri bildirimler feedbackler hakkında ne düşünüyorsun genel olarak?

Participant A6M: Ya bazı feedbacklerde mesela hocanın verdiği "aa" dedim "evet burayı unutmuşum nasıl unuturum" falan veya "nasıl yanlış yaparım" falan bazılarında böyle bir anlamadım sonradan "haa tamam" dedim. Yani faydalı oluyor feedbackler güzel gelişim sağladı bende.

Researcher: Yazma yeteneğini sence ne şekilde etkiliyor? Yani bu feedback dersen?

Participant A6M: Bu feedbackler mi?

Researcher: Hmmhmm, Evet.

Participant A6M: Yani hani bir değerlendirilmek güzel oluyor yani yazdığımız şeyi öyle öylece atmak değil de bir hani bir eleştiri almak bizi de pozitif etkiliyor. Hani gelişimimiz açısından.

Researcher: Peki sence etkin bir geri bildirim veya feedback işte nasıl olmalıdır? İdeal bir feedback senin için nedir, nasıl olur?

Participant A6M: Yani öncelikle gramer hatasıdır veya bir akışta hatadır bunlar olmalı bence veya içerikte hani konudan konuya sapmamak gerekir. Hani bunları da değerlendirirken bunları değerlendirerek feedback vermek bence çok daha faydalı olur.

Researcher: Anladım. Yani yazma ödevlerine peki geri bildirim almak sana genel olarak nasıl hissettin? Feedback, işte, geri bildirim hangisi senin içinde daha rahat oluyor? Geri bildirim mi? Feedback mi demem?

Participant A6M: Fark etmez.

Researcher: Tamam o zaman feedback diyeceğim. Yazma ödevlerine feedback almak nasıl hissettirdi sana?

Participant A6M: Ya dediğim gibi başta şaşırıyorum. Hani "bun nasıl kaçırırım" diye hani bir ödevi yaparken bazen hani hani acele yaptığım için kaçırdığım şeyler olabiliyor. Hani bazı şeyleri bazı verilen feedbackere şaşırıyordum ama bazıları da yani faydalı oluyor yani genel olarak.

Researcher: Tamam peki. Ödevlerin sınıfta öğretmen tarafından okunması veya düzeltilmesi sana genel olarak nasıl hissettirir?

Participant A6M: Ya güzel yani bence hani insanların önünde sınıfta okunması sınıfta değerlendirilmesi bence çok daha iyi. Çünkü hani birinin yaptığı yanlışı bir başkası da yapıyor çoğu zaman ve yani bir başkası oradan gördüğü zaman daha etkili olabiliyor yani ve hocanın da bizim yaptığımız ödevlere de hani değerlendirmesi bizim yaptığımız ödevlere değer verdiğini gösteriyor. Hani bu Biz öğrenciler için de biraz daha motive kaynağı olabiliyor hani.

Researcher: Evet anladım, doğru. Peki öğretim- öğretim değil de, Otomatik geri bildirim araçlarından yararlandın mı herhangi bir şekilde? Mesela grammarly, falan filan ChatGPT gibi.

Participant A6M: Yok, yok hiç.

Researcher: Yani burada dürüstlük önemli. Zaten hani şey önemli değil, yargılama falan herhangi bir şekilde yok yani.

Participant A6M: Yok yani nasıl işlediğini bile bilmiyorum.

Researcher: Hmm. Peki o zaman geri bildirimin eksik olduğu konular sence var mı varsa bunlar nelerdir?

Participant A6M: Yani. Hani bir deneme veya bilmiyorum geri bildirim bence her şeyle yapılır yani. Kusursuza- bir kusursuza yakın bir yazı yazabileceğimizi düşünmüyorum şu anda sınıfta. O yüzden bence geri bildirim daima olmalı yani.

Researcher: Peki eksik olduğunu düşündüğün konular olarak neler olabilir? Hani "şöyle yapılsa daha iyi olurdu, bu yoktu mesela" diyebileceğin.

Participant A6M: Yani hoca biraz daha gramere yönelmiş hani. Hani içeriğe de biraz şey yapsa olur. Gerçi onlarda da bir kaç tane şey atmış ama genelde gramer odaklı değerlendirmiş hoca, ki hazırlık döneminde biraz daha mantıklı olan bu aslında da hani lisans öğrencisi olmadığımız için. Ya içeri hakkında da biraz daha şey olsaydı iyi olabilirdi bence, daha iyi olabilirdi.

Researcher: Anladım. İşte bu genel olarak bu eksiklerin nasıl giderilebileceğini düşünüyorsun? Bahsettin ya mesela?

Participant A6M: Ya bu verilen feedbackleri dikkate alarak öğrencinin ve daha çok o düzeyde yoğunlaşarak daha çok yazılar yazarak daha çok yazılar okuyarak gelişeceğini düşünüyorum yani.

Researcher: Feedback'teki, mesela hocanın verdiği feedback'teki eksiklikleri mesela nasıl giderilebilir sence peki?

Participant A6M: Ya dediğim gibi hani konunun içeriğine bakılabilir biraz daha gramerdan ziyade ama o da hani hazırlık sürecinde olduğumuz için gramere bakmak aslında biraz daha mantıklı olabilir şu anda.

Researcher: Anladım, tabi. Peki aldığın geri bildirimleri yeterli buluyor musun, tam anlamıyla?

Participant A6M: Evet, yani gayet hani benim baktığımda hani okuduğumda feedbackleri tatmin ediyordu her zaman. Yani bir okuduğumda böyle "evet" diyordum. Hani "yanlış yapmışım". Evet bu "daha doğru olur böyle" diyordum.

Researcher: Peki "şu da olursa olurdu" falan diyeceğin, "bu yoktu mesela" diyebileceğin bir şey var mı?

Participant A6M: Yani "bu yoktu" dediğim değil de "nasıl bu hata oldu" falan dediğim vardı.

Researcher: Kendinle alakalı.

Participant A6M: Aynen yani bu sebeple ben noktalama işaretlerini düzeltmek- düzeltmesi falan işte hoca ben hani bulamıyordum, hani "nerede yanlış yapmışım" falan diye sonunda daha hani biraz daha incelersem bulabiliyorum anca onlar vardı bende.

Researcher: Anladım. Peki o zaman geri bildirim almak feedback almak, sendeki yazma isteğini etkiledi mi?

Participant A6M: Evet, yani dediğim gibi yazdığımız yazıya birinin değer vermesi hani birinin okuyup senin hatalarını söylemesi insanı daha öğrenciyi daha motive ediyor yani. Karşıdaki kişi bu yazıya benim azdığım yazılara değer veriyor ve ben de "daha özenli yazmalıyım, daha güzel yazmalıyım" hissini uyandırıyor yani.

Researcher: Yani. Peki bu yazma isteğini ne açılardan mesela etkiledi feedback almak?

Participant A6M: Ya dediğim gibi hani bu verilen feedbackler yani direkt motivasyon olarak etkiledi beni. Hani dediğim gibi "biraz daha özen göstermeliyim" "biraz daha değer vermeliyim" yazılara yani motivasyon açısından etkiliyor işte beni daha çok.

Researcher: Yani yazma motivasyonu artırıyor diyorsun.

Participant A6M: Evet.

Researcher: Peki yazma kaygısı hisseder misin veya anxiety şeklinde hani anxiety endişe gibi.

Participant A6M: Ya o kadar kaygı yok bende hiçbir zaman da dediğim gibi, yani bir mesela hoca hani bir anda derste işte şunun hakkında yazacak- yazacaksınız dediğinde bir ilk on saniye bir ne yazacağım acaba diye bir hisse kapılıyorum ama kafamda biraz yapınca hani dediğim gibi içerik çok önemli aslında, içerik hitap ediyorsa direkt hiç şey olmuyor mesela ya bir futbolla mesela ne bileyim orta çağdan yazı yazmak aynı şey olmuyor. O yüzden hani konu bana hitap etmezse biraz daha gerilebiliyorum ama çok da şey olmuyor.

Researcher: Peki bu kaygı olduğu zaman ne zamanlarda genelde oluyor? Gerçi az önce mesela dedin herhalde.

Participant A6M: Evet, yani dediğim gibi ilk ödev verildiği zaman bir kaygı oluşturuyor bence "acaba ne yapacağım" diye ama sonrasında halloluyor yani bir şekilde.

Researcher: Anladım. Peki geri bildirim almak, feedback almak senin kaygını yazma kaygını etkiliyor mu? Nasıl etkiliyor?

Participant A6M: Ya ben bu konularda çok kaygılı olmadığım için hani verilen feedbacklerde herhangi bir kaygı hissi uyandırmıyor veya hani bir endişemi azaltmıyor. Yani çok hani yazı yazarken writing dersinde falan çok endişe çok kaygılı ve çok stresli birisi değilim o yüzden daha rahatım yani.

Researcher: Anladım. Zaten kaygı yoksa sıkıntı yok o zaman. Genel olarak kendini yazmaya istekli hisseder misin?

Participant A6M: Evet ya bir şeyler yazmayı seviyorum ama hani bu şey yazılarını değil, yani hoca tahtaya bir şey yazar onu geçiririz onları çok sevmiyorum ama ne bileyim bir hani bir konu üzerinde bir şey yazma yani beni de geliştirip geliştirdiğini düşünüyorum bazen. Yani mesela hiç düşünmediğim bir şey bir anda yazı yazarken aklıma geliyor ve onu da yazıyorum "aa, evet" diyorum hani "bu da doğru" ve bu da aslında yani etkili oluyor benim için.

Researcher: Yani yazmayı genel olarak seviyor gibisin-

Participant A6M: Evet.

Researcher: -dediğini anlıyorum. Peki senin yazmaya istekli kılan durumlar neler?

Participant A6M: Ya dediğim gibi hani insan kendi başına yazı yazarken de keyif almalı ama bunlar birileri tarafından değerlendirildiğinde bu yazılar veya birileri tarafından pozitif veya negatif eleştiriler alınca bence daha değerli oluyor bir yazılar. Yoksa hani kendim yazıp bir yere astığımda kimse okumazsa o yazını sadece senin için değeri oluyor.

Researcher: Yani onun için değerlendirmesi.

Participant A6M: Evet yani bir feedback gelirse hoca tarafından dediğim gibi, bu o yazıyı çok daha değerli kılıyor bence.

Researcher: Anladım. Peki yazma motivasyonunun artırılması için daha başka neler yapılabilir?

Participant A6M: Ya başka, ya bazı insanlar bazı öğrenciler hani bu yazma konularında stresli kaygılı oluyorlar. Biraz endişeli oluyorlar ya böyle öğretmenler bence hani öğrencilere bu kaygıyı bu stresi hani azaltacak şeylerde tavsiyelerde bulunmalı bence yani "her gün yazı yaz" olur "yazılar oku" olur "nasıl yazı yazılır" falan bence böyle şeyler yani çok daha öğrenciler açısından etkili olabilir bence.

Researcher: Hmm, anladım. Peki bunun dışında yazma motivasyonunu en çok artan etkenler senin için nelerdir? Mesela neler olabilir? Konunun içeriği vesairedir veya işte-

Participant A6M: Dediğim gibi aynen içerik önemli hani değerlendirme geliştirmek, feedbackler bunlar önemli benim için. Ya bazen keyfi yazdığım yazılar da oluyor. Yani ama bu bence ya birileri okuması yazıyı bence çok daha önemli kılıyor yazıyı.

Researcher: Yani değerlendirilecek olması diyorsu.

Participant A6M: Evet.

Researcher: Yani mesela nede- ders şöyle ne zaman dersin "ben şöyle olunca daha fazla yazasım geliyor" dedirtecek bir durum ne olabilir, senin için var mıdır acaba?

Participant A6M: Yani, daha çok yazasım geliyor. Ya çok arkadaşım var, hani eğer ama az arkadaşım olsaydı biraz asosyal olsaydım yazmayı daha çok sevebilirdim. Hani insan yazarken kendi kendiyle de konuşuyor aslında-

Researcher: Evet.

Participant A6M: -bu yüzden hani daha etkili olabilir diye daha değerli olabilirdi writing veya yazma işi benim için ama hani sosyal birisi olduğum için çok böyle. Hani yazmak biraz daha yalnız işi-

Researcher: Sözlü- sözlü iletişimi daha çok kullanıyorsun?

Participant A6M: Aynen, aynen. Konuşarak yani çok yazıya dökmüyorum o yüzden.

Researcher: Yani, anladım. Peki aldığın geri bildirimde en çok yararlı olduğunu düşündüğün kısım veya kısımlar hangisi veya hangileri? Mesela işte gramer-yazım yanlışları düzeltmeleri veya işte içerik anlamında öneri ve yorumlar varsa aldıysan.

Participant A6M: Evet ya gramer hata- kendimden beklemediğim gramer hataları yapmıştım işte hoca yazdığında hani "gramer hatalarını düzelt" dediğinde bir şaşırmıştım. Hani "hata mı vardı?". Sonra tekrar baktığımda böyle 2-3 kez okuduğumda "evet" demiştim. Hani onlar benim için biraz şaşırtıcı olmuştu. Başka noktalama işaretleri falan onlara takılmıştım biraz ama onun dışında pek bir şey olmadı yani.

Researcher: Olmuyor. Peki o zaman yazılarının altında sağlanan yorumlar şeklinde bir şey vardı değil mi sizinkilerde? Hocanın eklediği yorumlar.

Participant A6M: He he yani ödevi teslim ettikten sonra hani ödevin altına yorum şeklinde veriyordu hoca feedbackleri.

Researcher: İşte bu- bu yorumlar genel olarak nasıl bir deneyim sağladı senin için, nasıl hissettirdi?

Participant A6M: Ya dediğim gibi kendimi geliştirdi biraz da hani bu yazılar bu verilen feedbackler beni biraz daha geliştirdiğini düşünüyorum.

Researcher: Peki his olarak nasıl bir his verdi seni- sana? Yani herhangi bir pozitif veya-

Participant A6M: Ya motivasyonumu arttırdı hani yazma motivasyonumu arttırdı diyebilirim direkt basit olarak.

Researcher: Öğretmenin sağladığı bu geri bildirimin etkinliğini nasıl artırabiliriz sence? Farklı neler yapılabilir?

Participant A6M: Dediğim gibi içeriğe biraz daha yönelebiliriz, ama o onda da yani hazırlık öğrencisi olduğumuz için hiç o bence gramere yönelmek şu an için daha mantıklı olabilir. Ama yani şu an pek aklıma bir şey gelmiyor onun için daha başka.

Researcher: Anladım. Peki o zaman, genel olarak böyleydi bu kadardı-

Appendix D: Sample ChatGPT Prompt Used For The Writing Tasks

You are an English writing evaluation and feedback assistant for instructors of entry-level EFL writing students. You provide feedback on students' writing assignments. Your tone is friendly, and you use simple words that are easy for beginner English learners to understand.

Please evaluate and provide feedback for the paragraph I will enter below.

First, write a 30–50-word overall evaluation for the paragraph, making friendly comments on its content and addressing the student. If possible, include informed comments about the student's subject (e.g., if they talk about a singer, mention some songs by the singer; if they talk about a computer game, mention some trivia about it as if you also experienced it).

Evaluate the paragraph on content, organization, vocabulary, language use, and mechanics:

Content: Is the content well-developed with necessary elements and a wide range of linking words? Are sentences and titles clearly defined? Is the text relevant to the descriptive paragraph genre? Is it focused and engaging? Are there repetitive or misused linking words? Are necessary details included, and are unnecessary details omitted? Are there elements that should be explained further or less?

Organization: Are words and idioms chosen effectively? Is the meaning clear, or does it read like a poor translation from Turkish? Warn against directly translated idioms or sayings from Turkish if detected.

Vocabulary: Are simple, compound, and complex sentences used effectively? Are there few errors in subject-verb agreement, verb tense, number (singular or plural), word order/function, articles, pronouns, and prepositions? Are there errors in negation and agreement? Does the writer understand sentence construction rules?

Mechanics: Are sentences well-constructed and clear in their aims? Are they too short or too long? Are there grammar, spelling, and overall mechanics errors?

Assess the paragraph through these elements without listing them in the feedback. Show where the student made mistakes or could improve, make suggestions, and provide specific examples from the text.

Make suggestions for improvement or areas where the writer can seek additional help from instructors or online resources. Explain some elements with SPECIFIC examples from the student's paragraph. If the text is shorter than 200 words, remind the student to write longer. Always boost the student's morale regardless of performance. The goal is to improve the student's writing skills and motivation. Use simple language and a friendly tone. Write the feedback as a paragraph about 100 words long and do not use lists or colons.

Appendix E: Consent Form for the Interviews

ÖĞRENCİLERİN OTOMATİK GERİBİLDİRİM ÜZERİNE TUTUMLARINA DAİR MÜLAKAT SORULARI

ONAM FORMU

Mehmet Akif Yıldırım tarafından yürütülen "İngilizce Yazma Derslerinde Otomatik Geri Bildirim ve Öğretmen Geri Bildiriminin Öğrenci Kaygı ve Motivasyonuna Etkisinin İncelenmesi" başlıklı araştırmaya davet edilmiş bulunuyorsunuz. Katılıp katılmamaya karar vermeden önce araştırmanın amacını ve yöntemlerini bilmeniz gerekmektedir. Sonuç olarak bu formu okumanız ve anlamanız önemlidir. Anlamadığınız veya sizin için açık olmayan herhangi bir konu olursa, sorabilirsiniz. Bu çalışmaya katılma kararı tamamen size aittir. Araştırma katılımını istediğiniz zaman reddetme ve katıldıktan sonra çalışmadan ayrılma seçeneğiniz vardır. Vereceğiniz yanıtlar, araştırmaya katılmayı kabul ettiğiniz anlamına gelecektir. Görüşme sırasında size sorulan soruları yanıtlarken, başkalarından gelecek herhangi bir baskı altında hissetmenize gerek yoktur. Çalışmanın kapsamına giren yanıtlarınız, daha sonra analiz edilebilmeleri için ses kaydına alınacaktır. Ses kayıtları sadece araştırma amacıyla kullanılacak olup, kimlik bilgileriniz gizli tutulacaktır.

Katılımınız için teşekkür ederim.

KATILIMCI İMZA

A GRUBU		B GRUBU	
1		1	
2		2	
3		3	
4		4	
5		5	
6		6	

Appendix F: Ethics Committee Report

E-82554930-050.01.04-355283
03971303

GİZLİ

T.C.
KARADENİZ TEKNİK ÜNİVERSİTESİ REKTÖRLÜĞÜ
Hukuk Müşavirliği

Sayı : E-82554930-050.01.04-355283 -771 22.03.2023
Konu : [redacted]ın Etik Kurul Kararı

İlgi : 21.03.2023 tarihli ve E-26014373-050.01.04-354831 sayılı yazı.

Sn. [redacted]

KTÜ-Sosyal Bilimler Enstitüsü Batı Dilleri ve Edebiyatı Yüksek Lisans Programı öğrencisi [redacted]ın "**İngilizce Yazma Derslerinde Otomatik Geri Bildirim ve Öğretmen Geri Bildiriminin Öğrenci Kaygı ve Motivasyonuna Etkisinin İncelenmesi**" adlı yüksek lisans tezi çalışmasında Etik Kurul incelemesi yapılmış ve başvurunuza onay verilmiştir.

Bilgilerinize ve gereğini rica ederim.

Prof. Dr. [redacted]
Rektör Yardımcısı

Ek: [redacted]'ın Etik Kurul Kararı

61080 – Trabzon / TÜRKİYE
Tel: 04623772107 Faks: 04623774399 www.ktu.edu.tr
Ayrıntılı Bilgi İçin İrtibat [redacted]
hukukmusavirligi@ktu.edu.tr

CURRICULUM VITAE

Mehmet Akif YILDIRIM graduated from Affan Anadolu High School in the same city. He then earned his degree in English Language and Literature from Karadeniz Technical University in the year 2020 as an honour student. He started his master's degree in the Applied Linguistics Program at Karadeniz Technical University in 2021 and is currently employed as a research assistant at the same department. His research interests include EFL writing instruction, technology in education, and language learning motivation.

YILDIRIM is single and speaks English, besides his mother tongue.