Journal of Educational Studies and Multidisciplinary Approaches (JESMA)

www.jesma.net

The Impact of Quillbot as an Automated Writing Evaluation Tool on EFL learners[1]

Nida Gürbüz[2]

To cite this article:

Gürbüz, N. (2024). The Impact of Quillbot as an Automated Writing Evaluation Tool on EFL learners. *Journal of Educational Studies and Multidisciplinary Approaches (JESMA)*, *4*(2),101-120. https://doi.org/10.51383/jesma.2024.101

[1] This article was presented at 15th International METU ELT Convention, which was hosted at Middle East Technical University, Ankara, Türkiye, on May 10-11, 2024.

[2] Lecturer, Necmettin Erbakan University, Faculty of Foreign Language Education, Department of Foreign Language Education, Konya, Türkiye, nida.unalan111@gmail.com

Journal of Educational Studies and Multidisciplinary Approaches (JESMA)
Volume 4, Issue 2 Year 2024 ISSN:2757-8747

The Impact of Quillbot as an Automated Writing Evaluation Tool on EFL learners

Nida GÜRBÜZ https://orcid.org/0009-0008-7110-9712

ARTICLE INFORMATION

Original Research

DOI: 10.51383/jesma.2024.101
Received 27 June 2024
Revision 14 July 2024
Accepted 05 September 2024

ABSTRACT

Automated Writing Evaluation (AWE) constitutes a machine-based approach employed in the evaluation of learners' writing proficiency. Recently, its widespread implementation has been geared towards enhancing learners' editing competencies. The primary objective of this research was to thoroughly assess QuillBot as an Artificial Intelligence (AI) tool system, specifically examining its utility for students in the context of paraphrasing and checking grammatical, mechanical and organizational mistakes. QuillBot stands as an online application designed to facilitate the paraphrasing of written content, mitigate the risk of plagiarism, condense lengthy sentences, and enhance grammatical precision to impart a professional appearance. This tool serves as a viable alternative for students confronted with the challenge of rephrasing content when manual efforts prove challenging or when innovative ideas for paraphrasing are not readily available. Employing a pre-post quasi-experimental research design with convenience sampling, the study focused on automated writing of a descriptive paragraph. A total of Turkish 48 English as a Foreign Language (EFL) learners were distributed across two groups, encompassing control group and experimental group. The quasi-experimental design incorporated triangulation with qualitative data derived from retrospective notes and reflective journal notes, while quantitative data underwent analysis via non-parametric tests and SPSS. Findings indicated a positive impact of the Quillbot on experimental group. The qualitative data analysis underscored participants' favourable assessments of both the software and the automated writing experiences.

Keywords: Online Application**,** Automated Writing Evaluation, Artificial Intelligence Machine-based approach, English as a Foreign Language (EFL)

Introduction

Over the past five decades, technology has undergone rapid development and has permeated nearly every facet of life, encompassing both industry and education. Its profound influence has played a pivotal role in the advancement of humanity and the evolution of international relations, fostering global connections among individuals from diverse geographical locations. The use of technology is transforming the way we think, learn, and carry out tasks in a ground-breaking manner, as revealed by (Collins & Halverson, 2010). Interactive educational technologies, including games, robots, virtual reality (VR), computer simulations, block-based computing, and internet-related tools, can enhance traditional ICT tools like projectors, digital whiteboards, and digital textbooks (Weng & Chiu, 2023).

The preceding century has experienced an unprecedented linear progression and innovation, primarily attributable to the substantial integration of technology into people's lives. This transformative integration has significantly altered human existence, with technological tools becoming ubiquitous in daily activities. Whereas conventional home telephones were once the norm, contemporary society predominantly employs mobile phones for a myriad of purposes. The advent of sophisticated technologies has led to the near-ubiquitous replacement of traditional home telephones by mobile phones. The swift and substantial changes witnessed by humanity within a relatively short span of time have also left an indelible impact on human behaviours and attitudes. Artificial Intelligence (AI) holds great potential in the realm of education by offering students individualized guidance, support, or feedback, and aiding educators.

The use of certain AI techniques could help achieve some or all of the following objectives, such as enhancing information retrieval, distinguishing intricate patterns, engaging in natural language conversation, and adapting based on "experience" (Peelle, 1971).

Several AI technologies contribute to language-related activities to enhance language quality and clarity, the sentence can be rephrased as follows: The system provides automated feedback for corrections and prompt responses, offers intelligent tutoring by answering questions and providing explanations, and personalizes learning experiences by delivering content tailored to individual needs. The extensive technical progress of AI encompasses several subcategories, including collective intelligence, computerised vision, common sense, data mining, and emotional intelligence. The subjects covered include game theory, image processing, natural language processing, neural networks, pattern recognition, and robotics.(Chong, 2020). AI systems utilize machine learning to execute specific tasks, such as providing feedback on learners' written work, translating written text, administering automated tests, or facilitating structured conversation practice through a chatbot application. The process of machine learning process involves utilizing statistical methods to analyze extensive sets of training data, recognize a particular way, construct a model (typically in the form of algorithms), and subsequently take actions based on that model (Hockly, 2023). Instructional Computer-Assisted Language Learning (ICALL) environments have been specifically crafted to enhance the learning journeys of students with varying language backgrounds and diverse learning objectives.

QuillBot is a widely used, easy-to-use and free paraphrasing AI tool as a machine learning system. As per Fitria (2021), the AI methodology integrates deep learning with specific natural language analysis approaches. Its automated function involves the extraction, addition, or alteration of words, thereby generating a novel sentence paraphrase, giving feedback and also providing an AI-based product that suggests paraphrases. When engaging the paraphrase function in QuillBot, the application undertakes the task of rewriting provided texts.

The importance of feedback in education has been well recognised due to its significant influence on students' learning and instructors' instructional habits and practices (Niu et al., 2021; Winstone & Carless, 2019). The significant influence of teachers' feedback in the classroom has frequently exerted considerable force on them, particularly in substantial classes and when it comes to assessment. The assessment practices of teachers are often considered to be primarily focused on "testing" rather than "assessment" when there is a lack of feedback (Wang et al., 2023).

Subject and Problem of the Research

This research has investigated the face-to-face evaluation problems both the challenges encountered by instructors and students during the process of learning and teaching languages. While numerous studies have examined these challenges, fewer have explored the role of automated writing evaluation tools in enhancing learning, writing, and instructional procedures. At nearly all educational institutions, computers are utilised for some tests and exams, such as proficiency exams, and TOEFL and also are used for preparing curriculum and timetable. Therefore, this study will explore the use of automated writing evaluation tools for improvements of students' writings while giving feedback. Karakaya illustrates that various visual elements such as images, animation, color, and visual design play a collaborative role with language in online communication. Platforms such as e-mail, instant messaging, chat rooms, Usenet groups, MOOs, blogs, and wikis facilitate novel modes of discourse, authorship, identity construction, and the establishment of learning communities and affinity groups that transcend international borders (Karakaya, 2010).

Research Questions

To fulfil the aim of this investigation, the subsequent research inquiries were developed:

1) How do Quillbot as an Artificial Intelligence Evaluation tool influence the quality and efficacy of students' writing skills and their confidence in writing capabilities of them?

2) Is there any advantage and effect of Quillbot as an Artificial Intelligence Evaluation tool on the enhancement of control and experimental group students' writing skills and educational outcomes by doing a comparison between the outcomes of the pre-test and post-test?

3) Is there a challenge in implementing Quillbot for the experimental group students?

4)Is there any effect of the application Padlet for on experimental group students' writings?

Purpose of the Study

This study aimed to evaluate the efficacy of QuillBot as an artificial intelligence tool for instructors and students in the domains of paraphrasing and grammar checking in English writing, specifically focusing on its free version. The study sought to examine the attitudes of EFL students and teachers towards automated evaluation tools and their use of technology in learning and teaching English. The study was carried out at a Turkish governmental institution, where the participants' writing levels and computer usage were assessed through a pre-test. The study also investigated the students' and teachers' attitudes towards QuillBot, an online evaluation tool. Qualitative, semi-structured face-to-face interviews (Bin Zou & Zhang, 2023) were used to examine the application of computer technology in teaching languages and to probe into the students' autonomy in using computer technology and the barriers to its implementation.

Significance of the Study

The importance of this study is underscored by its examination of English teachers' and students' perceptions of using automated writing evaluation tool and the illumination it provides on their utilization of technology in language teaching methodologies. Notably, there is a dearth of adequate data concerning the use of Quillbot, an online evaluation tool, and the attitudes of teachers and students toward this tool in Turkey. Consequently, this research stands as distinctive in furnishing comprehensive data concerning the impact of Quillbot on teachers and students in Turkey and their integration of computers in language learning and teaching methodologies.

REVIEW OF LITERATURE

The advancement of technology has significantly transformed the planet over time. The proliferation of technology worldwide has a profound impact on nearly every facet of life. The remarkable advancement of technology has significantly impacted education, making it one of the utmost crucial facets of life. The advent of technology has greatly enhanced the accessibility and effectiveness of education worldwide. Hence, technology has the potential to enhance student output, fluency, and potentially the writer's capacity to generate meaningful texts at an earlier stage in their language-learning curriculum. With a diverse range of tools and methods at hand, the integration of technology can offer the essential scaffolding and personalized support for students, mitigating certain challenges associated with acquiring proficiency in writing in a second language. The effectiveness of these technological elements in instruction should be assessed to optimize learners' time and efforts in the pursuit of second language acquisition (Raquel, 2004).

This need for assessment becomes even more relevant when considering the origins of artificial intelligence, which can be traced back 70 years to Alan Turing's 1950 publication, "Computing Machinery and Intelligence." In his seminal work, Turing posed the fundamental question of whether computers can exhibit thinking capabilities, and whether they can deceive humans into perceiving that they are engaging in conversation with a fellow human interlocutor rather than a computer. Understanding and leveraging these capabilities can significantly enhance the instructional process, making it more efficient and effective. The question is currently employed to evaluate the level of similarity between AI programmes and human people, referred to as the imitation game or Turing test. (Kwon et al., 2023). Artificial intelligence (AI) provides computer-assisted education with a previously deficient or completely absent attribute: the ability to engage in perceptive and "intelligent" interaction with a machine.

More precisely, the use of certain artificial intelligence approaches has the potential to advance one or more of the following objectives:

1) enhance the process of accessing information

2) accurately distinguish intricate patterns

3) employ a method to engage in discussion using natural language.

4) evolve and transform due to acquired 'knowledge'.

Within the educational context, these objectives can be transformed into benefits for the learner. For instance, artificial intelligence would significantly simplify the process, resulting in a broader range of feedback. This would enable a more interactive and intellectually stimulating partnership with the machine in an educational context, rather than viewing it as a mere taskmaster. Unfortunately, it is challenging for a tutor to fully meet the students' expectations due to their diverse abilities and categories in a classroom setting. Special education learners and advanced education learners have unique requirements. That is why, automated evaluation tools become indispensable. By using these tools, all students, regardless of their aptitude for learning, are given equal opportunities to acquire language skills at this stage. Intelligent systems demonstrate remarkable adaptability to the student's learning aptitudes and deficiencies.

Automated Writing Evaluation (AWE)

Al-Inbari and Al-Wasy mentions that Automated Writing Evaluation is a machine-based method employed to evaluate the writing of learners(Al-Inbari & Al-Wasy, 2023). Lately, there has been extensive use of this strategy to enhance learners' editing skills. As per Da Yan, the integration of technology into the process of acquiring linguistic skills has resulted in favourable outcomes on the proficiency and involvement of second language learners. Recent advancements, particularly the implementation of AI-based solutions, signify substantial strides made by educators in this domain. Advancements in technology accessibility and cost-effectiveness have heightened the adoption of technology-enhanced learning approaches in (L2) writing (Yan, 2023).

According to Zhang, the increasing interest in recent years has been focused on the utilization of automated writing evaluation (AWE) systems in (L2) writing. The majority of prior research focus on evaluating the impact of Automated Writing Evaluation (AWE) feedback on the writing of second language (L2) students. These studies generally examine aspects such as the scoring mechanism, validity, and reliability. However, they pay little attention to how L2 students actually interact with AWE input during the procedure of revising their writing (Zhang, 2020).

The initial implementation of automated essay scoring (AES) occurred in 1966. Subsequently, advancements in language processing technologies and statistical methods have facilitated the creation of new scoring engines, such as e-rater, Knowledge Analysis Technologies, and IntelliMetric. These engines are capable of analysing various aspects of text, including lexical, syntactic, semantic, and discourse features. Certain automated writing evaluation (AWE) instruments, like Criterion and MY Access, were created to offer formative feedback on diverse rhetorical and language-related aspects, alongside additional materials and automated assessments. While the majority of Automated Essay Scoring (AES) systems have been validated by demonstrating consistency and agreement with human raters, there is still a need to evaluate the educational and evaluative utility of Automated Writing Evaluation (AWE) (Li et al., 2015).

In Automated Essay Scoring, summative essay scores are usually generated by NLP computer programmes, which analyse the linguistic and semantic characteristics of the essays. AES algorithms are constructed using machine learning models that are trained on a collection of essays. These writings are evaluated by human raters using a scoring rubric. The models are then tested to guarantee that they can accurately assess a different set of essays. The algorithm's values are juxtaposed with their scores given by the raters to ensure their alignment and reliability. Undoubtedly, AES systems are often highly dependable, exhibiting significant agreement and associations with human raters. Automated scoring can provide several benefits that are similar to multiple-choice scoring, such as quick and constant scoring availability, reduced costs, and more consistent scoring. This can make it possible for some recommended testing programmes and learning settings to utilise more construct-relevant items, even if they were formerly too difficult to include (Ramineni et al., 2012). An increasing body of research is investigating the integration of automated writing evaluation (AWE) tools in writing classrooms, indicating a growing interest in their possible application for formative assessment. Nevertheless, like any assessment method, these tools require validation to ensure the accuracy of their intended interpretations and applications (Jim Ranalli & Chukharev-Hudilainen, 2017). Moreover, AWE can be used for summative assessment in writing instructions.

The creators define the feedback provided by the AWE tools MY Access, WriteToLearn, and Criterion as addressing both micro-level features of language, such as punctuation, spelling, mechanics, syntax, and use, as well as macro-level aspects.

Various AWE tools offer comprehensive and/or analytical evaluations to indicate the merits and shortcomings of a proposal. However, widely used AWE systems, such as Pigai, Quillbot, Criterion® and My Access!.

The purpose of these tools is to aid in writing for K-12 and college education, but they do not offer feedback on genre-specific writing or formative feedback on causal discourse. Some systems, such as

Grammarly mainly provides input on spelling, mechanics, and grammar. Conversely, Writing Pal provides feedback that is focused solely on writing processes and tactics. Criterion and similar systems provide feedback at both levels. (McCarthy et al., 2022).

Various forms and combinations of feedback are expected to provide different effects in terms of students' perceptions and more objective findings. In their study, Chen and Cheng found that automated feedback may be employed in the early phases of drafting and editing to provide formative assessment. This can then be combined with input from teachers and peers in the subsequent phases (Chen & Cheng, 2008). The significance of automated feedback, particularly the comprehensive score input provided by Automated Writing Evaluation (AWE). Link identified two further applications of the AWE score: as an initial evaluation or assessment before submitting a project or work for consideration to encourage student involvement in correction, and as a constituent of formal assessment in the classroom. Although educators as well as learners identified issues with automated scoring feedback, there was a lack of consistency between the AWE results and their evaluations.(Link et al., 2020).

There have been very few studies that have utilized this framework to assess the educational benefits of Automated Causal Discourse Evaluation Tool (ACDET) for language learners. The ACDET is designed to help ESL students improve their understanding of discourse activities that include discussing cause and effect relationships by evaluating two dimensions: Language Learning Potential and Meaning Focus. To evaluate the effectiveness of the Intelligent Academic Discourse Evaluator (IADE) tool in achieving comprehensive learning outcomes, observable learner behaviour was used as evidence. The evaluation was based on six criteria: Language learning potential, Learner fit, Meaning focus, Impact, Authenticity, and Practicality. The above standards give an in-depth perspective on the educational achievements that arise from the utilisation of IADE. (Feng, 2015).

There are several AWE tools such as; Criterion, Writing Pal, Pigai, Grammarly and Quillbot.

Criterion

Criterion provides a comprehensive evaluation that assesses the writing sample uploaded by users demonstrates a high level of overall quality as well as offering thorough diagnostic comments on mistakes across several areas. Domains may be obtained in a matter of seconds. It offers educators personalised reports that allow them to assess the success of their pupils. In addition, it offers students a full "Writer's Handbook" that encompasses input terms, interpretations of faults accompanied by instances, and recommendations for correcting errors and enhancing the organisation and development. (Dikli, 2006). Teachers have the ability to offer a prompt for writing derived from the Criterion subject library, which consists of over 400 writing prompts all types of writing and difficulty stages. Levels ranging from fourth grade to the higher education level. Teachers have the ability to provide feedback on students' written work by utilising the "Comment" or "Dialogue" sections.(Han & Sari, 2022.)

Writing Pal

W-Pal is an automated programme that evaluates writing and offers intelligent guidance for improving writing skills. It offers many methods for practicing writing, such as game-based strategy exercises and essay composition. The W-Pal training focuses on eight distinct components of writing: freewriting, planning, introduction development, body development, conclusion development, unity, paraphrasing and revision.(McCarthy et al., 2022).

Pigai

In addition to providing corrective feedback, it also includes holistic scoring, ranking, the top and lowest scores, and final remarks. The scoring model is calibrated using a comprehensive collection of texts that includes normal English, English essays written by students, and alternative English textbooks. Pigai computes a comprehensive score for each. The essay is evaluated by quantitatively comparing it

to other writings throughout its collection of texts in four specific domains: vocabulary, sentence structure, organisation, and subject relevancy (Zhang, 2020).

Grammarly

Grammarly's free edition offers input on spelling, punctuation, grammar, and conventions, such as space, capitalization, and spelling peculiar to different dialects. Grammarly promptly delivers comments for enhancement as soon as a document is posted online. The material that has been uploaded is shown to the left. The screen displays faults with red underlining, which is indirect feedback. The right half of the screen displays immediate feedback. Direct feedback provides detailed information about the exact type of issue, such as grammar, along with possible improvements and ideas (Koltovskaia, 2020).

Quillbot

Undoubtedly, we are familiar with the concepts of "paraphrase" or "rewrite" and now we will delve into their mechanics. The paraphrase tool operates by comparing two sources of text: the text that you type and the original text. One of them is called Quillbot.

QuillBot has undergone significant enhancements, introducing a range of features to augment its functionality. The initial addition was a summarization feature, facilitating the condensation of lengthy text into concise summaries. Subsequently, a grammar checker was integrated to assist users in identifying and rectifying grammatical errors in their written content. In more recent developments, QuillBot has expanded its capabilities by incorporating a citation generator, plagiarism checker, co-writer, and translator, thereby evolving into a comprehensive writing tool (Nour El Houda et al., 2023).

QuillBot is an online programme that can reword text, detect and prevent plagiarism, shorten long words, and improve grammar to achieve higher precision and a more sophisticated presentation. This research aims to assess the QuillBot as an Artificial Intelligence (AI) device that allows learners to paraphrase and rewrite texts in English. The QuillBot is accessible for both free and commercial versions. This study employs a descriptive qualitative methodology. The data used is an English abstract paper.

The findings indicate that QuillBot's text rephrasing tool employ multiple strategies for rephrasing the text, including: 1) employing equations or synonyms for paraphrasing, 2) altering the word's form for paraphrasing, 3) utilising active or passive sentence structures for paraphrasing, and 4) rearranging the word order in sentences for paraphrasing (Nguyen, 2023). This paraphrase employs the Standard Mode, which aims to maintain a balance in modifying the text while ensuring that the original meaning remains intact. Fitria states that the AI method combines deep learning with other techniques for natural language processing. Its function is to automatically remove, insert, or modify words in order to generate a completely new phrase (Nur Fitria, 2021).

The complimentary edition of QuillBot permits paraphrasing of content within a limit of 400 characters. The advanced iteration of this tool extends the character limit to a maximum of 10,000. QuillBot encompasses seven noteworthy features, among them is the Standard Mode, which guarantees that alterations made to the input text maintain its intended meaning while augmenting its originality.

Fluency Mode is designed to prioritize the natural flow of writing and uphold precise and suitable English grammar. It introduces subtle adjustments to the text while maintaining its original meaning effectively.

On the other hand, Creativity Mode places a strong emphasis on substantial modifications to the provided text. However, such extensive alterations may result in a shift in meaning or overall coherence. This mode proves beneficial when the objective is to make the text distinctly different from the original content.

Furthermore, the Creative+ Mode allows for more intuitive and extensive modifications, especially concerning common phrases or expressions.

The Shorten Mode is designed to compress the information while preserving its intended significance, making it especially valuable for minimising word count or total text length. On the other hand, the Expand Mode aims to lengthen the text by including additional words, which might be beneficial when trying to boost the overall word count.

Finally, the Formal Mode is dedicated to altering the text to be more appropriate for a formal audience. This approach proves highly effective for writing in academic or corporate contexts.

Using Padlet in Classrooms

Padlet facilitates the instantaneous display of various file formats on the wall, offering an invaluable method for organizing teaching materials. Both the teacher and students populated the respective Padlets with videos, documents, and images utilized during the lessons. This approach enabled students to conveniently access and download resources directly from the platform. A class diary is a document generated promptly after each session, encompassing a session summary along with qualitative data pertaining to the teaching process. Padlet facilitates interactive debates among students by providing various collaboration options. These options include the ability to add comments to others' posts or to engage in reactions using a voting system (Beltrán-Martín, 2019).

The results of this study indicate that the utilisation of Padlet software has the potential to enhance students' learning by enhancing their level of involvement in activities both within and beyond the classroom. The programme offers extensive functionality and enables students to work alone or collaboratively on individual and group assignments (Megat et al., 2020). In this study, Padlet serves as an electronic portfolio, allowing students to archive their written works. It provides a platform for collaborative learning, enabling students to access and review each other's writings. The researcher utilizes Padlet to provide feedback and assign grades to the students. Additionally, before an examination, the researcher downloads the students' writings and distributes them to their respective groups, facilitating collaborative study and revision of both individual and peer contributions.

THEORETICAL FRAMEWORK

Theoretical frameworks lead, formulate, and explain how research studies are conducted. The following hypothesis has shaped the researcher’s approach to defining, explaining, and understanding Turkish students' preparatory classroom writing and technology use.
Early CALL classrooms just had a computer that only the teacher could utilise. This phase of CALL is known as ‘Behaviouristic CALL’ (Tafazoli & Golshan, 2014). Computers were mostly used for repeated workouts and skill reinforcement. Structural CALL is associated with the 1970s and 1980s. This period's technology resembled classroom methods, according to some. Grammar-translation and audio-lingual methods dominated Structural CALL language courses. Tafazoli and Golshan defined the next step of Computer-Assisted Language Learning as 'Communicative CALL' (2014). In the 1980s and 1990s, systems prioritised accuracy and fluency equally instead of correctness. The main activities emphasised communication skills in English education.

Warschauer & Healey (1998) call the newest step of Computer-Assisted Language Learning (CALL) 'Integrative CALL'. Technology in the 21st-century language classroom is the focus of this phase. It emphasises agency and the learner's independence. The focus is on interpersonal communication, with technology used to access real materials showing how the language is used. Integrative CALL requires technological instruments to be incorporated into language study rather than used sometimes in computer labs.Tafazoli & Golshan (2014)
Bax uses 'Restricted CALL,' 'Open CALL,' and 'Integrated CALL.' Limitated CALL emphasises language system content with tasks like rewriting texts and answering closed-ended questions (Bax,

2003). This period involves little student interaction. Open CALL, on the other hand, emphasises computer-mediated communication and linguistic abilities. Simulation games like Second Life help students use tools. Finally, Integrated CALL integrates language skills and employs classroom-appropriate technologies. This method encourages student-technology engagement. (Erol Aslıhan Nur, 2022). Stockwell stresses that CALL environment substantially affects technology selection and use (Stockwell, 2012a). Many CALL researchers have also noted the importance of context (Wang, Lili, 2023). The phenomena is complex, involving the school setting, policies, instructional technique, educational objectives, students' backgrounds, expectations, prerequisites, and other factors.

METHODOLOGY

Research Design

This study utilised a quasi-experimental research approach known as a pre-test and post-test control group design. The researcher designated intact classes as the experimental group and the control group.

The researcher gathered data of quantitative and qualitative kind in order to address the study inquiries. The study's quantitative component analysed the influence of integrated automated-teacher input regarding students' analytic composition grades and error count. This was done by comparing the results of pre-test and post-test writing assignments. The qualitative data collected from student comments provided insights into students' perspectives on obtaining automated feedback on their writing proficiency.

The study gathered qualitative data through individual interviews with teachers and students to analyse their use of Quillbot and assess the perceived advantages and challenges associated with utilising the system's corrective feedback. Simultaneously with the collection of this data, quantitative data was also collected. This included the list of submissions for each paper as recorded by the researcher, as well as AWE error reports. The post-test underwent evaluation by two instructors, thereby enabling the assessment of interrater reliability. The purpose of collecting this data was to analyse the impact of AWE for evaluating and providing comments on students' writing practice to enhance the correctness of their paper drafts.

The Research Procedures

The researcher taught both the experimental and control groups. All of them were given identical guidance on academic and introspective writing during their lesson. In order to accomplish the objectives of this study, the subsequent methods were implemented:

1. A total of 48 students were divided into two separate groups: a control group comprising 24 individuals and an experimental group consisting of 24 individuals. Prior to beginning the experiment, each group was instructed to compose a descriptive paragraph. This activity served as a pre-test to assess writing proficiency both of the groups.

2. The paragraphs in the pre-test were evaluated based on a rubric created by the KET Assessment of Writing Scale by Cambridge English. The rubric for grading paragraphs has five aspects, each assigned a certain percentage. These factors include content and organisation. The user's text mentions the topics of vocabulary, language use, and mechanics, with a numerical reference to 10 and a vocabulary level of 4. The number is 6.

3. A dedicated classroom was established to provide comprehensive instruction to the two members of the experimental group on the software (Quillbot). The instruction covered topics such as software features, usage guidelines, handling programme feedback, and exploring various functions.

4. The participants in the experimental group were instructed to submit their colleagues' paragraphs to the Quillbot programme, while they themselves were prompted to submit their own paragraphs. Subsequently, the experimental group was required to obtain the programme feedback,

implement all essential revisions, and ultimately transcribe the final versions of the altered paragraphs exactly as they are shown in the programme.

5. The experimental group was presented with the retrospective note question during this step.

6. The essays from both the control and experimental groups were evaluated as a post-test of the treatment. The evaluation was done using rubrics by Cambridge KET.

7. The feedback technique for the control group only consisted of written instructor input pertaining to the structure and substance of the work. The teacher offered criticism on the form, namely about grammar, use, and mechanics, by emphasising the paragraphs of the pupils' contained words, phrases, or sentences that required correction or improvement. The comments lacked clarity and was not straightforward. The teacher was permitted to utilise her pre-existing collection of mistake codes during the research to provide an authentic feedback process. The mistake codes and the feedback given by the teacher were analysed via Quillbot.

Participants

The research was conducted at a Turkish state university, located in Konya in 2023-2024 Fall Term. A non-probability self-selection strategy was employed to enlist 48 participants who were English as a Foreign Language (EFL) students currently enrolled in a preparatory class. Students were evaluated based on the institution's English Language Admission Test and determined to have a minimum English proficiency level of Common European Framework of Reference (CEFR) A2. The study adhered to established ethical protocols, with subjects providing voluntary consent to participate. During the discussion with the participants, it was clearly specified that their participation in the study would not be compensated with any extra academic credit.

The study included a cohort of 48 English as a Foreign Language (EFL) students, consisting of 20 males and 28 females, with ages ranging from 20 to 30 years old. The participants were categorised into two distinct groups: a control group who received feedback on their assignments from a human teacher, and an experimental group which received feedback created by the Quillbot programme.

Data Collection Tools

Pre- test and post-test writing tasks

For both pre-test and post-test, the students were required to write a 150–200-word length descriptive paragraph on a singular subject that had been chosen for all students. Both tasks were selected from the paragraph topics of the book" Language Hub Main course" and the writing course assumed that assignments were similar in terms of the students' degree of competence, experiences of writing classes, educational interests, and cultural features, with regards to topic familiarity. The students were allowed 40 minutes to write their paragraphs using Padlet application which had been introduced to students beforehand. The course teacher accepted the paragraphs through Padlet software to ensure that every student could write. The students composed their essays on the subsequent subjects, which were not previously deliberated with them.

Pre-test writing topics

The topics that students wrote "Describe your neighbourhood? Describe your best friend? The happiest (most frightened, worst etc.) moment in your life. Support your position with explanations and examples from your own experiences, observations or reading."

Post-test writing topics

The topics that were written by students at the end of the implementation "Describe your dream apartment? Describe your first day of school? Describe a cinema that you watched recently? Describe your favourite film? Give examples and explanations for your answer.

Student Reflections

The students in the experimental group were instructed to write their comments on the process of completing their tasks under the combined automated-teacher feedback condition. They composed their views after the conclusion of the treatment, once they had acquired sufficient expertise in the feedback process they had been subjected to. The researchers devised a set of questions to direct the students in articulating their encounters with and perspectives on getting automated responses from Quillbot. Students also gave reflection about using Padlet. The pupils were permitted to compose their views in Turkish as it was anticipated that they would articulate their opinions more effectively in their own language. The researcher of the study translated the quotations provided in this article into English. What is more, following the implementation of the tool, the students were interviewed by the researcher. Each of the 24 students in the experimental group was interviewed in semi-structured interviews lasting 30 minutes. The interviews took place in secluded offices and were recorded using a digital voice recorder. The recordings were then transcribed word for word.

Data Analysis

The statistical data analysis in this study was conducted using SPSS Version 22 for Windows. The study also incorporated the data derived from the retrospective notes and semi-structured interviews. In order to ensure the dependability of the coding, the paragraphs from the pre- and post-tests were assessed by the two researchers and an additional colleague, utilising the rubric provided by Cambridge KET as described before. The third research question was addressed by analysing the quantitative data using ANOVA. The purpose of the test was to ascertain whether there were notable disparities between the writing grades of experimental and control groups. The interview data collected from students were subjected to qualitative analysis, during which emergent categories were discovered.

Subsequently, conclusions were derived from these categories in order to address the second study inquiry.

ETHICAL CONSIDERATION

The study adhered to established ethical protocols, with subjects providing voluntary consent to participate. During the discussion with the participants, it was clearly specified that their participation in the study would not be compensated with any extra academic credit.

The data was collected from students who were allowed to express their perspectives in Turkish, since it was expected that they would communicate their viewpoints more efficiently in their own tongue. The researcher of the study rendered the quotations presented in this article into English. In addition, the researcher administered semi-structured interview questions to the pupils. She captured the students' vocalisations and subsequently translated them into English. Students underwent individual semi-structured interviews, with each session spanning a duration of 30 minutes. The interviews were conducted in private offices and recorded using a digital voice recorder. The audio recordings were then transcribed verbatim.

Ethical review board name: Necmettin Erbakan University
Date of ethics review decision: 02/02/2024
Ethics assessment document issue number: 17798

FINDINGS

Upon examining the research questions;

RQ1. How do Quillbot as an Artificial Intelligence Evaluation tool influence the quality and efficacy of students' writing skills and their confidence in writing capabilities of them?

The examination of the data obtained from the student reflections yielded two overarching themes, namely good impressions and negative perceptions. Each category is outlined below. Excerpts of exemplary quotations extracted from the reflection papers.

The favorable perceptions of the experimental group of students towards automated writing evaluation tools were evident in their positive views of the automated feedback they received in a writing course. Initially the students discovered that automatic feedback was beneficial for improving their writing skills. According to them, the most advantageous aspect of automated feedback is its promptness and availability to students, provided they have an internet connection. The individuals held the belief that receiving prompt feedback facilitated the process of rectifying their mistakes, particularly when their thoughts and essay structures were still vivid in their brains. Furthermore, they claimed that promptly rectifying mistakes improved long-lasting retention of knowledge. The following remarks, offered by one of the participants, clarified these points:

"One aspect of automatic response that I particularly appreciated was its immediate accessibility upon submitting my paragraph. I strongly feel that committing errors is really advantageous for the process of learning. Additionally, promptly recognising and acknowledging these blunders proves to be exceedingly useful."

"Efficient acquisition of knowledge. Furthermore, rectifying my mistakes was facilitated by the fact that the ideas I attempted to convey in my paragraph remained vivid in my memory. Upon receiving criticism on my errors, I am compelled to contemplate the intended message of my writing, which subsequently induces fatigue and disinclination towards error correction."

The students expressed their enjoyment of automated feedback due to their appreciation for incorporating technology into their writing course. According to their assessment, writing a paragraph on the computer was more convenient and effortless. Superior to the act of writing with pen and paper in terms of enjoyment. In addition, the participants acknowledged that obtaining prompt feedback on their written work was a source of motivation for them. A participant expressed the following comment:

"Receiving a comprehensive score from the programme was highly motivating for me." On every occasion, I exerted maximum effort to generate an improved paragraph with the intention of achieving the highest possible score, which is 9. Furthermore, when rectifying my mistakes and resubmitting my paragraph, I observed an instantaneous surge in my score, which motivated me to rectify all my problems and produce flawless paragraphs for future tasks."

The students reported experiencing fewer mistakes and improvements in their writing process. Some specifically mentioned:

"The frequency of errors in my writing diminished, and my writing process became more efficient. The generation of words and sentences became more fluid and expeditious. In the initial assessment, I produced less content and encountered numerous punctuations and spelling errors. However, in the subsequent evaluation, when I scrutinized my writing, I observed a reduction in mistakes, particularly in punctuation and spelling. Engaging in extensive paragraph writing during lessons and receiving feedback through QuillBot proved to be highly advantageous for my overall writing proficiency."

RQ2- Is there any advantage and effect of Quillbot as an Artificial Intelligence Evaluation tool on the enhancement of control and experimental group students' writing skills and educational outcomes by comparing the results of the pre-test and post-test?

To examine the influence of automated-teacher feedback on students' analytic writing scores, the pre-test and post-test writing scores of both the experimental and control groups were compared.

Table 1-The table shows two-factor ANOVA analysis of the students' writing scores, the pre-test and post-test averages and standard deviations.

	Pre-test		Post-test	
	M	SD	M	SD
Experimental	11,88	2,112	16,67	3,031
Control	14,21	2,904	13.75	2,863

Based on the data shown in the table, the pre-test scores of the experimental group students were 11.88, while their post-test scores were 16.67. Upon analysing the pre-test and post-test results of the pupils in the control group, they were found to be 14.21 and 13.75, respectively.
Statistics regarding students' English writing exam scores are shown in the table 2 below, categorized by variables of score, test, and group.

Table 2: Tests of Between-Subjects Effects

Dependent Variable: PRE-TEST_POST-TEST

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	280.083[a]	3	93.361	12.333	.000	.287
Intercept	19153.500	1	19153.500	2530.270	.000	.965
GROUP	112.667	1	112.667	14.884	.000	.139
TEST	2.042	1	2.042	.270	.605	.003
GROUP * TEST	165.375	1	165.375	21.847	.000	.192
Error	696.417	92	7.570			
Total	20130.000	96				
Corrected Total	976.500	95				

a. R Squared = .287 (Adjusted R Squared = .264)
0,01-0.09 (small), 0,09-0,25 (medium) and >0,25 (large)

The joint effect of the Test and Group variables on students' scores was found to be significant, $F(1,92)=21.847$, $p<.05$, Partial $\eta^2=0.192$. In other words, being in the control or experimental group has a different impact on increasing the final test scores. The increase in writing final test scores for students in the experimental group was statistically significantly higher compared to students in the control group and it was found statistically significant as $p<.05$.

RQ3: Is there a challenge in implementing an Quillbot for the experimental group students?

Upon comparing the two categories, it becomes evident that the students' unfavourable impressions are far less. The primary aspect of Quillbot that received the most criticism from students was its inability to offer content input.

Additionally, they asserted that the warning messages are challenging to comprehend since they do not explicitly indicate the accurate format of the problem. A participant articulated his or her viewpoint in the following manner:

"Automated feedback solely displays and classifies our errors, but it would be more advantageous if it also offers the correct format." I encountered challenges in identifying the appropriate format of the mistakes that I made due to my inability to comprehend the warning messages.''

The students noted that the comments provided about the organisation and growth aspect was ambiguous. Quillbot utilises colour codes to visually represent the structural components of the essay that has been submitted, including the thesis statement, key ideas, supporting sentences, and ending phrases. Pupils conveyed that the advising messages displayed while clicking the colour codes are redundant and lacking in useful information. A participant expressed his opinion on this matter with the following statement:

"Upon each submission of my paragraph, I consistently received advisory messages regarding the development and organisation of my essay." I believe that these signals lack specificity and fail to provide me with comprehensive information on the organisational components in my paragraph.''

In addition, the participants reported that the programme occasionally produces inaccurate error messages and high level of translation. In relation to this topic, the students provided the accompanying illustrations:

"The programme indicates a grammatical error in the highlighted portion of the sentence, however, there is actually no issue present. Furthermore, I believe that this programme is translating a text at an advanced level, but I need a more basic translation. Translating according to our proficiency level would be beneficial."

"Quillbot issued a repetition warning despite the fact that I used the word just a limited number of times."

A student expressed dissatisfaction with QuillBot due to its limitations on mobile devices. The student mentioned:

" *The application sometimes overlooks errors when used on a mobile phone, prompting the need to switch to a computer for more accurate corrections. The absence of detailed explanations for corrections on the mobile version was highlighted as a drawback, contributing to the perception that studying on a computer is a more efficient alternative."*

RQ4- Is there any effect of the application Padlet for experimental group students' results?

One of the students clarified some points about Padlet:

"The conventional method of submitting handwritten papers to the teacher has been replaced by the utilization of Padlet, allowing our teacher to evaluate our paragraphs at her convenience. This digital approach eliminates the need for paper, promoting environmental sustainability. Subsequently, we engage with QuillBot to refine and enhance our writing, streamlining the overall process."

Another student expressed an opinion about Padlet;

"Through the platform of Padlet, I can engage in collaborative efforts with my peers. This collaborative space facilitates the sharing of writing outcomes and feedback among friends. In preparation for exams, I have the opportunity to review and analyze the results and mistakes of my friends, fostering a collective and supportive learning environment."

DISCUSSION AND RESULT

As per the findings of Turgay Han and Elif Sarı, the study demonstrated that the writing scores of the students in the experimental group showed improvement and reduced their mistakes following the combined automated-teacher feedback (Han & Sari, 2022.). This study assessed the writing accomplishments of the experimental group students and juxtaposed it with the control group, which received comprehensive evaluation from the instructor. The results indicated that there was no statistically significant disparity between the two groups in terms of the overall score they obtained from the analytic scoring scale.

There may be variations in the feedback conditions used in the previous research compared to the feedback circumstances used in the current investigation. Different approaches of obtaining input may uncover conflicting outcomes. The disparity can be explained by the students' reflections. They held the conviction that automated feedback was more effective for learning as it allowed them to promptly address their shortcomings.

In addition, the AWE system offered personalised assistance at any time and location. Providing students with the opportunity to rectify their mistakes before to submitting their assignments for teacher evaluation might have enhanced their autonomy, fostering a learner-centred atmosphere that facilitated independent learning. Furthermore, the disparity between the two groups in terms of the enhancement in grammar and mechanics might also be attributed to the fact that Quillbot identified a far greater number of errors compared to the teacher, therefore heightening the students' consciousness of their own flaws. While the teacher has the skills to offer thorough criticism similar to Quillbot, it is possible that she overlooked certain errors due to exhaustion. The majority of participants held the belief that automated feedback significantly enhanced their overall writing proficiency. The participants conveyed that immediate feedback motivated them to complete assignments promptly, driven by their curiosity about their performance. Additionally, they were motivated to revise their compositions further as their grades improved upon addressing the problematic areas. Furthermore, rather of grading several papers on a daily basis, EFL instructors may allocate their time and effort towards seeking out methods to enhance the efficacy of their writing teaching for students.

It is important to acknowledge that the success of AWE integration relies on several aspects, including students' proficiency with technology and instructors' preparedness and eagerness to include technology into their teaching.

Stephanie Link, Mohaddeseh Mehrzad, and Mohammad Rahimi suggest that future research in AWE (Automated Writing Evaluation) should incorporate consistent and standardised measurements of the writing construct(Link et al., 2022) . This would enable meaningful comparisons across different studies and enhance the overall value of the research. The revelation of AWE feedback can be enhanced. What is lacking in this analysis is a criterion, or a group of criteria, that assists in evaluating constructs that are of significant importance to SLA scholars, such as genre knowledge, structure, and idea growth.

The implications of this study's findings should be considered in view of various constraints. Initially, the current investigation failed to manage potential factors that may have affected the pupils' independent utilisation.

Sarıcaoğlu and Bilki mention that regarding AWE, it is important to consider factors such as error correction backgrounds, motivation levels, assignment assessment standards, and other related aspects. The study lacks information regarding the quantity of feedback given by teachers on students' English language in their written assignments and whether the linguistic quality influenced the grading of these assignments (Saricaoglu & Bilki, 2021). The importance that students place on the English language in their assignments could be influenced by the importance given to language by their professors, potentially affecting the students' utilization of Automated Writing Evaluation (AWE). Future research endeavours may benefit from including the measurement of these potential parameters.

Furthermore, although the results indicate that the instructor's involvement is crucial for students to effectively utilise AWE, it remains uncertain whether specific teacher tactics and styles were most helpful in enhancing students' writing with AWE.

Based on the available evidence, it is clear that teachers who hold a favourable view of incorporating innovative approaches in their lessons are more inclined to invest additional effort to make their instruction more impactful. Establishing examples and delivering lectures on the utilization of computer technologies is of paramount significance. Moreover, it is crucial to allow prospective teachers to experiment with these tools, spend time familiarizing themselves with them, and contemplate how they can integrate them into their lessons. Consequently, this experience will prompt them to adopt a positive mindset and enthusiastically embrace the adoption of new technological resources for use within the classroom (Ateş Özdemir, 2013).

LIMITATIONS AND IMPLICATIONS

A limitation of this research is the experimental focus on solely 48 participants within a preparatory class. Expanding the experiment to include a larger and more diverse sample, involving students from various universities, could enhance the generalizability and comparability of the results.

Future research might employ other metrics of accuracy and explore the impact of AWE corrective techniques. Feedback can enhance students' cognitive and metalinguistic development and its relationship to independent learning.

The present study possesses several constraints that must be taken into account while evaluating its results. In the present study, the participants consisted of motivated students enrolled in English Preparatory Classes. Open to receiving any sort of critique that would assist in enhancing their writing proficiency. Subsequent research may be undertaken on non-preparatory English lessons and the outcomes can be juxtaposed with those gained from all academic disciplines. Furthermore, due to the study being conducted in an instructional context, it was not feasible to utilise a genuine experimental research design. Utilising arbitrary choice and participants’ assignment may lead to more dependable outcomes.

Further study should explore the possibility of augmenting the participant pool and examining the impact of AWEs) on advanced learners. Furthermore, future research should investigate the methodologies that teachers can employ to facilitate peer-review activities. Expanding the range of qualitative data, such as conducting classroom observations and interviews can be advantageous for evaluating the effectiveness of an AWE programme. One further constraint of this study is the absence of questionnaire implementation, potentially impacting the participants' replies. Hence, it would be

more advantageous for further investigations to employ an anonymous questionnaire in order to guarantee the dependability of the survey.

REFERENCES

Al-Inbari, F. A. Y., & Al-Wasy, B. Q. M. (2023). The impact of automated writing evaluation (AWE) on EFL learners' peer and self-editing. *Education and Information Technologies*, *28*(6), 6645–6665. https://doi.org/10.1007/s10639-022-11458-x

Ateş Özdemir, E. (2013). *IMPROVING PRESERVICE ENGLISH LANGUAGE TEACHERS' ICT SKILLS & DEVELOPING POSITIVE ATTITUDES TOWARDS THE USE OF TECHNOLOGY IN LANGUAGE TEACHING THROUGH AN ONLINE COURSE*.

Bax, S. (2003). CALL—past, present and future. *System*, *31*(1), 13–28.

Beltrán-Martín, I. (2019). *Using Padlet for collaborative learning*. 1–8. https://doi.org/10.4995/HEAD19.2019.9188

Bin Zou, Z. L., Qinglang Lyu, Yining Han, & Zhang, W. (2023). Exploring students' acceptance of an artificial intelligence speech evaluation program for EFL speaking practice: An application of the Integrated Model of Technology Acceptance. *Computer Assisted Language Learning*, *0*(0), 1–26. https://doi.org/10.1080/09588221.2023.2278608

Chen, C.-F. E., & Cheng, W.-Y. E. C. (2008). *Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes*.

Chong, J. V. V. (2020). *Perspectives on Artificial Intelligence in Education: A Study of Public Elementary School Teachers* [PhD Thesis]. Biola University.

Collins, A., & Halverson, R. (2010). Technology supports for lifelong learning. *International Encyclopedia of Education*, 184–188.

Dikli, S. (2006). An overview of automated scoring of essays. *The Journal of Technology, Learning and Assessment*, *5*(1).

Erol Aslıhan Nur. (n.d.). *A SUGGESTED SYLLABUS FOR TECHNOLOGY ENHANCED LANGUAGE TEACHING CLASS AT ELT DEPARTMENTS*.

Feng, H.-H. (2015). *Designing, implementing, and evaluating an automated writing evaluation tool for improving EFL graduate students' abstract writing: A case in Taiwan*.

Han, T., & Sari, E. (n.d.). An investigation on the use of automated feedback in Turkish EFL students' writing classes. *Computer Assisted Language Learning*, 1–24. https://doi.org/10.1080/09588221.2022.2067179

Hockly, N. (2023). Artificial Intelligence in English Language Teaching: The Good, the Bad and the Ugly. *RELC Journal*, *54*(2), 445–451. https://doi.org/10.1177/00336882231168504

Jim Ranalli, S. L., & Chukharev-Hudilainen, E. (2017). Automated writing evaluation for formative assessment of second language writing: Investigating the accuracy and usefulness of feedback as part of argument-based validation. *Educational Psychology*, *37*(1), 8–25. https://doi.org/10.1080/01443410.2015.1136407

Karakaya, Kadir. (2010). *Taffa*. Middle East Technical University.

Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. *Assessing Writing*, *44*, 100450. https://doi.org/10.1016/j.asw.2020.100450

Kwon, S. K., Shin, D., & Lee, Y. (2023). *The application of chatbot as an L2 writing practice tool*.

Li, J., Link, S., & Hegelheimer, V. (2015). Rethinking the role of automated writing evaluation (AWE) feedback in ESL writing instruction. *Journal of Second Language Writing*, *27*, 1–18. https://doi.org/10.1016/j.jslw.2014.10.004

Link, S., Mehrzad, M., & Rahimi, M. (2020). Impact of automated writing evaluation on teacher feedback, student revision, and writing improvement. *Computer Assisted Language Learning*, *35*, 1–30. https://doi.org/10.1080/09588221.2020.1743323

Link, S., Mehrzad, M., & Rahimi, M. (2022). Impact of automated writing evaluation on teacher feedback, student revision, and writing improvement. *Computer Assisted Language Learning*, *35*(4), 605–634. https://doi.org/10.1080/09588221.2020.1743323

McCarthy, K. S., Roscoe, R. D., Allen, L. K., Likens, A. D., & McNamara, D. S. (2022). Automated writing evaluation: Does spelling and grammar feedback support high-quality writing and revision? *Assessing Writing*, *52*, 100608. https://doi.org/10.1016/j.asw.2022.100608

Megat, N., Mohd Azmi, N., Yusoff, R., Shariff, S., & Wan Hassan, W. A. (2020). Enhancing Classroom Engagement Through Padlet as a Learning Tool: A Case Study. *International Journal of Innovative Computing*, *10*. https://doi.org/10.11113/ijic.v10n1.250

Nguyen, X. (2023). *Using the Online Paraphrasing Tool Quillbot to Assist Students in Paraphrasing the Source Information: English-majored Students' Perceptions*. 21–27. https://doi.org/10.21467/proceedings.150.3
Niu, Z., Zhong, G., & Yu, H. (2021). A review on the attention mechanism of deep learning. *Neurocomputing*, *452*, 48–62.

Nour El Houda, C., ZERHOUNI, A., & NAILI, R. (2023). *The Impact of Analysing QuillBot Products on Enhancing EFL Learners' Paraphrasing Skills* [PhD Thesis].

Nur Fitria, T. (2021). QuillBot as an online tool: Students' alternative in paraphrasing and rewriting of English writing. *Englisia Journal*, *9*, 183. https://doi.org/10.22373/ej.v9i1.10233

PEELLE, H. A. . (1971). *A STUDY OF COMPUTER-ASSISTED LEARNING WITH ARTIFICIAL INTELLIGENCE GAMES.*

Ramineni, C., Trapani, C. S., Williamson, D. M., Davey, T., & Bridgeman, B. (2012). EVALUATION OF THE E-RATER® SCORING ENGINE FOR THE TOEFL® INDEPENDENT AND INTEGRATED PROMPTS. *ETS Research Report Series*, *2012*(1), i–51. https://doi.org/10.1002/j.2333-8504.2012.tb02288.x

Raquel, M. N. (2004). *EFFECTS OF TECHNOLOGY-ENHANCED LANGUAGE LEARNING ON SECOND LANGUAGE COMPOSITION OF UNIVERSITY-LEVEL INTERMEDIATE SPANISH STUDENTS*.

Saricaoglu, A., & Bilki, Z. (2021). Voluntary use of automated writing evaluation by content course students. *ReCALL*, *33*(3), 265–277. https://doi.org/10.1017/S0958344021000021

Stockwell, G. (2012). *Computer-assisted language learning: Diversity in research and practice*. Cambridge University Press.

Tafazoli, D., & Golshan, N. (2014). Review of computer-assisted language learning: History, merits & barriers. *International Journal of Language and Linguistics*, *2*, 32–38. https://doi.org/10.11648/j.ijll.s.2014020501.15

Wang, Lili. (n.d.). *A Mixed-Methods Study on the Technology Use of Chinese EFL Teachers in High School Classrooms. (Under the direction of Dr. Hiller Spires).*

Wang, Y., Derakhshan, A., Pan, Z., & Ghiasvand, F. (2023). Chinese EFL Teachers' Writing Assessment Feedback Literacy: A Scale Development and Validation Study. *Assessing Writing*, *56*, 100726. https://doi.org/10.1016/j.asw.2023.100726

Warschauer, M., & Healey, D. (1998). Computers and language learning: An overview. *Language Teaching*, *31*(2), 57–71.

Weng, X., & Chiu, T. K. F. (2023). Instructional design and learning outcomes of intelligent computer assisted language learning: Systematic review in the field. *Computers and Education: Artificial Intelligence*, *4*, 100117. https://doi.org/10.1016/j.caeai.2022.100117

Winstone, N., & Carless, D. (2019). *Designing effective feedback processes in higher education: A learning-focused approach*. Routledge.

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, *28*(11), 13943–13967. https://doi.org/10.1007/s10639-023-11742-4

Zhang, Z. (Victor). (2020). Engaging with automated writing evaluation (AWE) feedback on L2 writing: Student perceptions and revisions. *Assessing Writing*, *43*, 100439. https://doi.org/10.1016/j.asw.2019.100439

APPENDIX-1: SEMI-STRUCTURED INTERVIEW QUESTIONS

1. Do you think Padlet is useful in improving your writing skills?
2. Are you willing to try to use AI tools such as Quillbot for academic English learning in the future?
3. What do you think of the automated evaluation system of Quillbot? Do you think it provides consistent scores with the human graders?
4. What do you think of the feedback and error analysis system of this tool? Any suggestions?
5. What were the most interesting things/advantages in this implementation?
6. What were the most difficult things/disadvantages in this implementation?
7. Compared to classroom, did you like writing on Padlet? What was the difference between feedback on a paper and feedback on the application?
8. Do you want to add more information about this implementation?

APPENDIX-2: WRITING RUBRIC BY KET

B1 Preliminary Writing Examiners use the following assessment scale, extracted from the one on the next page:

B1	Content	Communicative Achievement	Organisation	Language
5	All content is relevant to the task. Target reader is fully informed	Uses the conventions of the communicative task to hold the target reader's attention and communicate straightforward ideas.	Text is generally well organised and coherent, using a variety of linking words and cohesive devices.	Uses a range of everyday vocabulary appropriately, with occasional inappropriate use of less common lexis. Uses a range of simple and some complex grammatical forms with a good degree of control. Errors do not impede communication.
4	*Performance shares features of Bands 3 and 5.*			
3	Minor irrelevances and/or omissions may be present. Target reader is on the whole informed.	Uses the conventions of the communicative task in generally appropriate ways to communicate straightforward ideas.	Text is connected and coherent, using basic linking words and a limited number of cohesive devices.	Uses everyday vocabulary generally appropriately, while occasionally overusing certain lexis. Uses simple grammatical forms with a good degree of control. While errors are noticeable, meaning can still be determined.
2	*Performance shares features of Bands 1 and 3.*			
1	Irrelevances and misinterpretation of task may be present. Target reader is minimally informed.	Produces text that communicates simple ideas in simple ways.	Text is connected using basic, high-frequency linking words.	Uses basic vocabulary reasonably appropriately. Uses simple grammatical forms with some degree of control. Errors may impede meaning at times.
0	Content is totally irrelevant. Target reader is not informed.	*Performance below Band 1.*		

Biographical notes

Nida Gürbüz is a lecturer within the Department of Preparatory Class at Necmettin Erbakan University, Konya. Concurrently, she pursues doctoral studies ELT department. Her scholarly endeavors focus on technology and academic writing particularly the integration of artificial intelligence into language learning and teaching methodologies.