


Automated Writing Evaluation in EFL Contexts: A Review of Effectiveness, Impact, and Pedagogical Implications

Tahani Aldosemani, Prince Sattam bin Abdulaziz University, Saudi Arabia*

 <https://orcid.org/0000-0002-2347-1564>

Hussein Assalahi, King Abdulaziz University, Saudi Arabia

Areej Lhothali, King Abdulaziz University, Saudi Arabia

Maram Albsisi, King Abdulaziz University, Saudi Arabia

ABSTRACT

This paper explores the literature on AWE feedback, particularly its perceived impact on enhancing EFL student writing proficiency. Prior research highlighted the contribution of AWE in fostering learner autonomy and alleviating teacher workloads, with a substantial focus on student engagement with AWE feedback. This review strives to illuminate these facets and offer critical insights on AWE effectiveness, feedback quality, reliability, and usefulness. Guided by the research questions, 16 studies were selected, adopting specific inclusion criteria to assess the effectiveness of AWE in enhancing EFL learner writing performance. Recommendations and implications from the reviewed articles regarding AWE implementation were synthesized and discussed. The review concludes that AWE can improve EFL student writing skills, with varying effectiveness based on student proficiency levels. AWE provides quality feedback and can be a reliable and valuable tool. However, despite its effectiveness, human intervention is essential to maximize its outcomes and mitigate limitations.

KEYWORDS

Automated Writing Evaluation, Classroom Strategies, English as a Foreign Language, Feedback Effectiveness, Feedback Quality, Reliability, Usefulness

INTRODUCTION

Automated writing evaluation (AWE), also known as *automated writing feedback* (AWF), is a computerized feedback system that uses natural language processing (NLP) to analyze and provide feedback on student writing regarding grammar, style, and content. These systems aim to enhance writing proficiency by delivering immediate and customized feedback on language errors applicable

DOI: 10.4018/IJCALLT.329962

*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

across various writing tasks. Recently, there has been a significant surge in interest concerning the pedagogical aspects of AWE feedback within English as a foreign language (EFL) writing research. The direction of this research has markedly shifted from considering AWE as an alternative to human evaluators to examining how AWE-generated feedback can enhance the quality of EFL writing (Ranalli et al., 2017; Hibert, 2019).

The potential of AI-powered AWE and automated text scoring (ATS) systems in enhancing writing in English as a second language has been explored in numerous studies (Chen & Cheng, 2008; Grimes & Warschauer, 2010; Alikaniotis et al., 2016; Tang & Wu, 2017; Zhang, 2021). The literature yielded several key areas encompassing various aspects of AWE use in EFL classroom settings. First, despite their effectiveness in improving writing skills and providing quick and measurable feedback, AWE tools exhibit varying levels of effectiveness. User attitudes towards these tools are also mixed, with an appreciation for their convenience and immediacy but criticism for their inability to provide nuanced feedback. In addition, while AWE systems emphasize the role of revision and practice, the constant submission of redrafts does not necessarily lead to improved writing skills. Therefore, the necessity for human intervention and implementation strategies persists, reinforcing that these tools can supplement the writing activity without replacing the role of instructor feedback.

BACKGROUND

Effectiveness and Limitations of AWE

Research indicates a clear potential for AWE tools to enhance the writing process and provide valuable feedback (Dikli, 2006). These tools, underpinned by theoretical foundations, offer opportunities for deliberate practice and holistic feedback on student writing, with or without instructor support. AWE potential for EFLs is reflected in studies such as Lee (2020), in which students using AWE feedback over an extended period reported improved writing competence. Despite its potential to enhance the writing process, AWE tools exhibit varying levels of effectiveness and have limitations that need to be considered (Hibert, 2019). While competent in detecting lexical errors, AWE feedback may not accurately evaluate university student essays, such as those offered by Pigai (<http://en.pigai.org/>; an AWE system popular in China that uses NLP to assess and provides feedback on English language essays; Gao, 2021).

The effectiveness of AWE can vary significantly depending on student proficiency levels. Xu and Zhang (2021) found that lower-level students benefited more from Pigai AWE feedback than their more proficient peers. On the contrary, Huang and Renandya (2020) found that integrating AWE did not result in improved drafts for less proficient students, indicating that more targeted approaches may be needed based on the varying needs of different learner profiles. Studies underscore the importance of these tools to provide more targeted, specific feedback that addresses individual student needs (El-Ebyary & Windeatt, 2010; Chapelle et al., 2015; Koltovskaia, 2020). Wang, Chen, and Cheng (2006) noted that while AWE was beneficial in improving certain formal aspects of writing, it fell short in identifying more complex issues of coherence and idea development. Considering their potential advantages and limitations, this highlights a critical balance in the perceived effectiveness of such tools.

User Perceptions and Attitudes Towards AWE

Previous studies shed light on mixed perceptions of AWE among users. Zhang (2021) highlighted that student attitudes significantly influence their usage of AWE for revisions. Some users appreciate its immediacy, convenience, and ability to identify specific errors. However, others expressed dissatisfaction with the feedback's lack of specificity and applicability to broader aspects of writing, such as coherence and content development (Lee, 2020). Indeed, student perceptions of AWE and willingness to use such systems vary widely, as reported in different studies. Zhai and Ma (2021) and Huang and Renandya (2020) discussed student willingness to use AWE and their satisfaction

with the software. Positive attitudes are typically attributed to the system's effectiveness in detecting errors and the immediacy and convenience of feedback. On the other hand, negative perceptions are often linked to the feedback's lack of specificity and helpful writing tips. Bai and Hu (2017) similarly reported mixed views about AWE. A high uptake of suggestions was needed, particularly concerning mechanical errors, grammar, and collocations, suggesting varying trust in the feedback provided. The literature varies on the perceptions of AWE feedback reliability, with some highlighting a degree of skepticism (Bai & Hu, 2017) and others expressing satisfaction with the feedback's accuracy (Li et al., 2015; Ranalli et al., 2017). The diverse perspectives indicate a need for a more conclusive examination of the quality and usefulness of AWE feedback.

The Role of Practice and Instructor in AWE Activities

As facilitated by AWE systems, revision, and deliberate practice are critical factors in improving writing skills. Cognitive models of writing (Warschauer & Grimes, 2008; Stevenson & Phakiti, 2013) emphasize the importance of revision in writing improvement. Through AWE systems, students can practice writing, revise their work, and even submit multiple drafts based on automated feedback, leading to improved drafts and better writing skills. In addition, several studies still highlight the need for human feedback (Li et al., 2015). For instance, AWE feedback was often perceived as more beneficial when combined with teacher feedback, addressing challenges in interpreting input and providing detailed feedback (Boud & Molloy, 2013). Similarly, Mohsen (2019) noted that while AWE improved students' writing skills to an acceptable level, human intervention was still needed to ensure the accuracy of AWE programs and address issues like clarity, coherence, and ambiguity in student writing.

Indeed, the success of AWE tools in EFL hinges upon their effectiveness and applicability, often manifesting in a dichotomy between the tools' potential and their shortcomings. Therefore, this study aims to shed light on AWE feedback from various perspectives, encompassing its effectiveness, feedback quality, reliability, and usefulness. It also delves into how to strategize an optimal implementation of AWE tools. By doing so, this study aspires to contribute to the evolving discourse on AWE technology, recognizing its promise while emphasizing the necessity for continuous development and improvement.

The Current Study

Various studies have highlighted the effectiveness of AWE in identifying errors and offering feedback that improves writing tasks (Lee, 2020; Xu & Zhang, 2021). However, the nuances of AWE's effectiveness may depend on several factors underpinning the necessity for further exploration and comprehension. Therefore, this study probes into the AWE feedback quality, reliability, and usefulness aspects to comprehensively evaluate AWE tools. Lastly, considering the evident impact of AWE on EFL learners' writing performance, it is crucial to identify the most effective strategies for its implementation in classrooms. The potential for AWE tools to support teachers and students is evident, though literature suggests the need for careful integration balanced with human intervention (Boud & Molloy, 2013; Mohsen, 2019). This study aims to investigate and highlight the advantages, limitations, and best practices in integrating AWE feedback in EFL classrooms in the literature, thus maximizing the benefits of this innovative technology in second-language education. Given these goals, the primary research questions this study is addressing are as follows:

- Q1:** How effective is AWE feedback in improving EFL writing performance?
- Q2:** What are the common aspects of AWE feedback quality, reliability, and usefulness?
- Q3:** What are the most recommended strategies for implementing AWE feedback in EFL classrooms?

The research questions focus on AWE feedback in EFL classrooms. The first question investigates the effectiveness of AWE feedback, which directly evaluates its primary function in enhancing writing

skills. The second question delves into AWE feedback quality, reliability, and usefulness. These aspects, crucial for any educational tool, directly impact its acceptance by teachers and students and contribute to its overall effectiveness (Gao, 2021). The third question seeks to identify the best strategies for implementing AWE feedback. The effectiveness, quality, reliability, and usefulness of AWE feedback can be significantly influenced by its implementation. Therefore, effective strategies can maximize AWE feedback benefits and mitigate potential drawbacks. In essence, these questions are linked by the common thread of understanding and optimizing the use of AWE feedback in EFL education. They each tackle a different facet of the issue, providing a comprehensive view that can inform better EFL teaching and learning practices.

METHODOLOGY

Literature Search and Inclusion Criteria

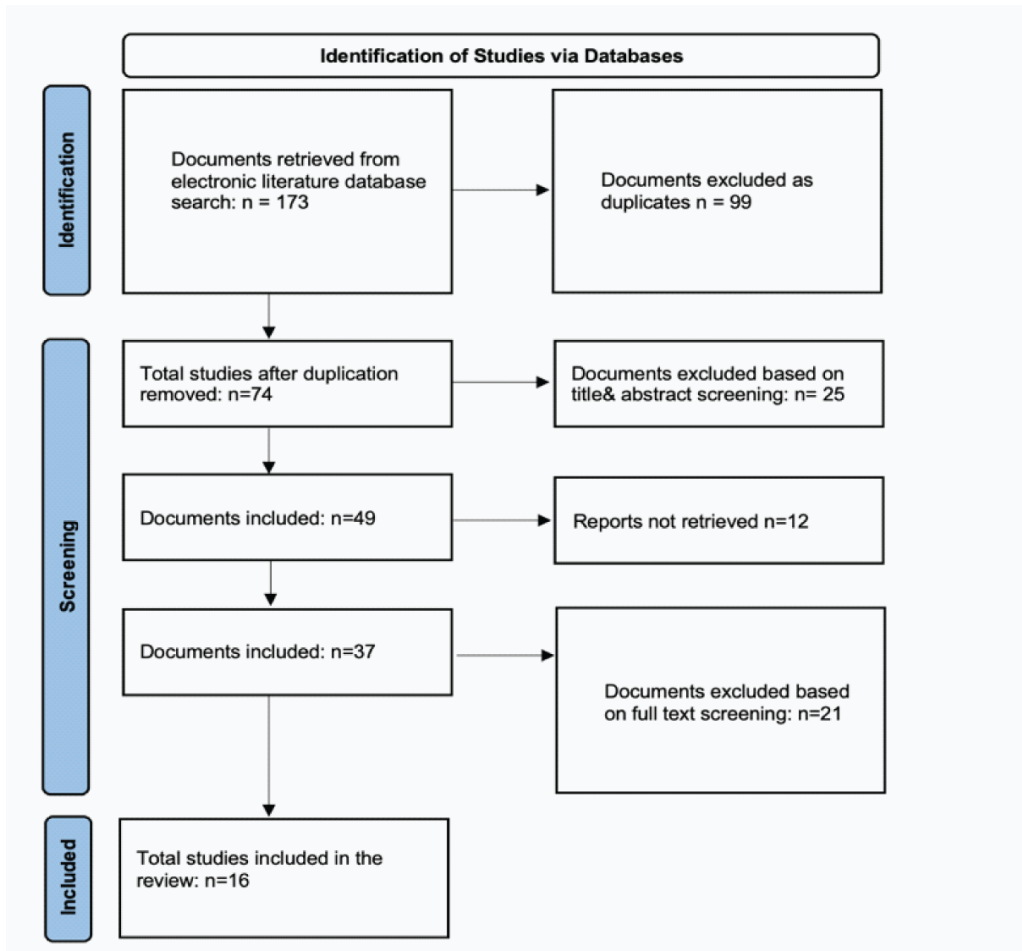
The first step in identifying the body of research included in this review was to perform a database search for articles published in the last ten years. The first stage was to conduct an advanced search in Scopus, ScienceDirect, Sage Journals, ProQuest, OBESCO, Oxford Academic Journals, and Google Scholar for articles about AWE feedback in EFL and ESL contexts. These resources cover a vast range of subjects and disciplines; their database typically hosts peer-reviewed research, which adds to their reliability and credibility and is continuously updated with new research.

The search yielded 173 research articles. The inclusion criterion for screening the pooled articles involved three main parameters. First was empirical research on AWE, because feedback is provided to EFL/ESL students based on observed and measured phenomena that derive knowledge from experience rather than from theory or belief. This criterion was included to ensure that the research examined real-world applications of AWE and its effects on EFL/ESL learners. This approach also guarantees that the research provides data that can be measured and quantified, leading to more robust conclusions. Second was to be conducted in higher education settings. This criterion was chosen because higher education institutions often have more resources to implement AWE technologies. Students in these settings may have different needs and experiences than younger students or those in informal learning environments. It also ensures a certain level of academic rigor and professionalism in the studies included. Third was to be published between 2010 and 2022. This criterion ensures that the studies included are relatively recent and reflect the current state of AWE technologies. This period saw significant advancements in AWE tools and their application, so studies from this time will likely provide the most relevant and up-to-date insights.

The exclusion criteria for this study were based on excluding any articles that do not meet all of the above inclusion criteria. Furthermore, articles that are not in English, do not provide sufficient data for analysis, or have significant methodological flaws were also excluded. These criteria aim to ensure high research quality, relevancy, and applicability. It is crucial to have a clear and rigorous set of criteria to ensure the research included is reliable, valid, and can contribute to a comprehensive understanding of the use of AWE in EFL/ESL contexts.

A first round of examining the abstracts based on the criteria reduced the number to 74 articles. Further examination was conducted based on the criteria, resulting in 49 research articles for review. A pre-final review based on a detailed reading of the articles resulted in 37 research journals. A final review based on full-text screening was included, focusing on 16 articles for this review (see Figure 1). All searches were carried out using the following terms: English as a second language (ESL), EFL, second language (L2), or English for specific purposes (ESP), AWE, and AWF. While this research included articles about using AWE in EFL contexts, some used ESL interchangeably. This research included articles about using AWE in both EFL and ESL contexts. While the two terms are often used interchangeably, they differ. EFL generally refers to learning English in non-English-speaking countries, while ESL is often associated with learning English in English-speaking countries.

Figure 1. Inclusion and exclusion process



Rationale

In the current AWE research, most studies have focused primarily on the validity of scoring models, aiming to measure how closely AWE can replicate scores assigned by human raters (Shermis et al., 2008). In addition, numerous studies focused on the potential of AWE tools to alleviate the workload associated with human feedback. As such, AWE tools' pedagogical applications are often underrepresented in the literature (Cheng & Zhang, 2021), contributing to a general hesitance towards incorporating these systems into writing instruction (Wilson & Roscoe, 2019). Therefore, this research critically explores the effectiveness of AWE in EFL classrooms in L2 writing enhancement, focusing on the quality, reliability, and usefulness of the automated feedback students and instructors receive. This highlights the importance of the research questions guiding this review, which aim to scrutinize AWE feedback's effectiveness and implementation strategies more closely and underscore the need for comprehensive investigations into these areas.

RESULTS AND DISCUSSION

To answer research questions guiding this review, 16 studies have been selected following this study's methodology and criteria to examine AWE's effectiveness in improving writing performance for EFL learners. Four studies were analyzed to answer the first question (see Table 1) and 12 studies to answer the second research question (see Tables 2–4), and for the third research question, aggregated recommendations and implications on AWE implementation from reviewed articles are discussed and synthesized.

Effectiveness in Improving EFL Writing Performance

The research literature underscores the positive impact of AWE feedback on EFL writing proficiency. Cheng and Zhang (2021) conducted an in-depth analysis of EFL student writing skills. The study's findings revealed AWE's considerable influence on student writing proficiency, as evidenced by improved accuracy, fluency, content, and structural quality. The value of AWE was also confirmed, as students acknowledged its usefulness in refining their EFL writing, particularly in terms of fluency and accuracy. The researchers stressed the necessity of follow-up rewriting activities after AWE, as these tools offer students opportunities to apply and solidify the knowledge gained from revision. They confirmed that AWE aids students in achieving a well-rounded development in second-language writing. However, their study recommended a balanced approach to teaching in implementing AWE, discouraging teachers from focusing excessively on linguistic errors while neglecting issues related to content and structure during feedback provision. In a related study, Tang and Rich (2017) explored the effectiveness of AWE using various methods to collect evidence on the tool's effectiveness. Their findings illustrated that AWE stimulated student interaction and motivation, encouraging continuous essay writing and revision. They concluded that AWE is a powerful tool for teachers in EFL classrooms, promoting student-teacher engagement in the writing process. It also facilitated dynamic writing assessment and improved student autonomy, increasing student engagement in the writing tasks.

Li, Feng, and Saricaoglu (2017) adopted a comprehensive mixed-method approach to explore students' perceptions and experiences of the effectiveness of AWE. The findings demonstrated that AWE feedback significantly decreased the number of errors in student writing, with the most impactful feedback relating to the grammatical aspects of their compositions. Furthermore, the study concluded that such feedback was practical and could boost human-computer interaction, foster immediate improvements, and result in drafts with substantially reduced errors. These benefits were possible as the initial exposure to corrective feedback prompted students to correct their errors immediately. However, the study also noted some potential drawbacks of AWE. It pointed out that AWE could inadvertently result in minimal long-term improvements and foster an over-reliance on AWE, causing students to pay less attention to grammatical accuracy in their subsequent drafts. Consequently, the researchers concluded that while AWE could be instrumental in reducing most errors in short-term writing activities, its benefits may be less substantial for long-term error reduction.

Li et al. (2015) employed a mixed-method approach to examine data from multiple data resources. The principal findings suggest that following AWE feedback, EFLs demonstrated an increased engagement in writing and improved linguistic writing accuracy and error types. However, the study also revealed that instructional methods could influence these outcomes. For instance, implementing a minimum score as a requirement for submission could impact results. The study concluded that it is crucial to consider the instructors' pedagogical methods and perspectives to optimize the benefits of AWE and effectively address potential issues during its use. This approach ensures that students derive the maximum possible advantage from the AWE system.

Several key takeaways emerged from these studies on AWE effectiveness in EFL writing. First, all discussed studies show that AWE can significantly enhance student writing performance, particularly in accuracy, fluency, content quality, and structural organization. Second, AWE feedback, especially grammar, was beneficial and actionable. Students appreciated its value and found it helpful

in improving their writing. Third, AWE feedback effectively reduced errors in student writing in the short term. It facilitated immediate correction of errors and produced drafts with fewer mistakes. In addition, AWE was found to increase student interaction and motivation, thus encouraging continuous writing and revision.

Further, AWE tools fostered increased autonomy and independence among students, thus allowing them to be more actively involved in their writing tasks. The teaching methods and implementation strategies influenced the effectiveness of AWE. In addition, optimal benefits from AWE were strongly linked to effective pedagogical practices. Further, to address potential issues and maximize the benefits of AWE, careful consideration of pedagogical methods was recommended.

Finally, despite its immediate benefits, studies showed that AWE could result in limited long-term gains. There was a risk of students becoming over-reliant on the system, potentially leading to less attention to grammatical accuracy in future drafts. In summary, while AWE provides immediate benefits in improving EFL student writing, its use should be thoughtfully integrated into teaching practices for long-term learning. Further, its usage needs to be supplemented with traditional feedback and correction mechanisms to ensure the comprehensive development of writing skills. Table 1 summarizes AWE's effectiveness, advantages, and limitations in EFL writing performance.

Common Aspects of AWE Feedback Quality, Reliability, and Usefulness

Prior research has underscored the importance of the quality, reliability, and practical usefulness of feedback given by AWE systems. These characteristics are pivotal for any pedagogical tool, as they

Table 1. AWE feedback effectiveness in improving EFL learner writing performance (Li et al., 2017; Tang & Rich, 2017; Cheng & Zhang, 2021)

AWE Advantages for EFL Classrooms		AWE Limitations for EFL Classrooms	
Key Advantages	Description	Key Limitations	Description
Improvement in writing skills	AWE has a substantial impact on improving writing skills, ranging from fluency and accuracy to content quality and organizational structure. Significant reduction in student writing errors, particularly in the grammar component.	Lack of long-term gains	AWE offers fewer benefits for error reduction in the long term. Students may need to fully internalize and apply error corrections in the future.
Engagement and motivation	AWE enhances student motivation and engagement during the writing and revision processes. AWE is an interactive platform encouraging student involvement and motivating their interest in writing tasks for continuous learning and improvement.	Reliance and decreased attention to grammar	Over-reliance on AWE might lead to less attention to grammatical accuracy in future drafts. This could affect the development of the critical skills needed for self-editing and revision.
Students autonomy and self-directed learning skills	AWE fosters student autonomy in writing tasks, allowing them to take ownership of their learning. It enhances student involvement in the writing tasks and be more attentive to their grammar, form, and usage, improving their self-directed learning skills.	Dependent on instructional approaches	AWE benefits significantly depend on the instructional approaches used and might be less effective across all teaching methodologies.
Teaching efficiency and pedagogical versatility	AWE supports balanced L2 writing development, offering dynamic assessment capabilities that can be tailored to suit various instructional approaches. It frees up teachers' time, allowing them to focus more on content instruction and personalized student support. It can handle high volumes of student work, offering scalability for larger classrooms	Overemphasis on linguistic errors	Although this limitation is related to implementation rather than an inherent limitation, teachers should avoid focusing on linguistic errors at the expense of content and organization issues. Teachers should balance correcting errors and content organization to maintain writing coherence.

directly influence its adoption and acceptance by both teachers and students, thereby significantly determining its overall effectiveness (Gao, 2021). Each of these aspects is examined in detail.

AWE Feedback Quality

Kerr (2020) stated that feedback quality, essential to EFL writing, involves active learner participation and adherence to writing class rules through varied feedback types (Wulandari, 2022). Foster (2019) used semi-structured interviews to study student perceptions, revealing that AWE feedback aided essay drafting and content structure, fostering better essays, improving self-confidence, speeding up writing, and reducing drafts. The study suggests that AWE can streamline assignment reviews and increase software familiarity. Allen, Likens, and Mcnamara (2018) reported that AWE showed minor overall essay revisions, providing low-level feedback without substantially impacting essay quality. The study underlined the significance of AWE training for EFL writing.

Gao (2021) evaluated AWE feedback quality by analyzing 104 student essays, revealing its accuracy in identifying errors consistent with professional teachers, particularly lexical ones. However, it missed some linguistic errors. While students praised its lexical feedback, they criticized its syntax and collocation error detection. The study suggests using AWE as a support to other feedback methods. Saricaoglu and Bilki (2021) assessed AWE's impact on student improvement over two assignments in two courses, showing AWE feedback reduced error rates across multiple grammar categories, with significant student writing improvement and error reduction.

The key insights from the referenced studies on AWE quality show that feedback quality in EFL writing is crucial. It helps writers structure their ideas effectively and align them with specific learning goals. The feedback enhanced learners' active involvement and supported their adherence to the acceptable rules of an L2 writing class (Kerr, 2020). Further, AWE feedback is particularly beneficial during the drafting phase of essay writing. It can guide structuring content, enhancing student self-confidence, accelerating the writing process, and reducing the number of drafts (Foster, 2020).

In addition, these studies show the importance of training students to use AWE for EFL writing. Although lower-level feedback on spelling and grammar did not significantly improve essay quality (Allen et al., 2018); however, AWE was adept at identifying errors and offered valuable feedback on lexical errors and suggestions for improvement. Overall, AWE leads to substantial improvements in students writing. The research found a significant error reduction across several categories, including grammar, usage, writing mechanics, and writing style (Saricaoglu & Bilki, 2021). Despite its benefits, AWE should be used sparingly and in conjunction with other feedback mechanisms, acknowledging its limitations in detecting certain error types (Gao, 2021). In summary, while AWE can enhance the drafting process, improve feedback quality, and reduce errors, training students in its usage is crucial. Its limitations should also be recognized, and its use should be combined with other feedback mechanisms to support students comprehensively. Table 2 summarizes the key advantages and limitations of AWE feedback quality.

AWE Reliability

Past research has measured the reliability of AWE software through its error detection precision and recall accuracy percentages, aiding quantitative and qualitative evaluations (Li et al., 2015). Precision pertains to the ratio of detected errors to total AWE-detected errors, while recall accuracy refers to correctly detected errors to those identified by expert raters (Burstein et al., 2004). Precision rates vary across error types, representing only a fraction of total text errors. Studies have quantitatively examined students' trust in AWE feedback, considering factors like the editing process, drafts, revisions, and error types (Bai & Hu, 2017). Other studies used surveys and interviews to capture student perceptions of AWE reliability.

Bai and Hu (2017) used a mixed-method approach to assess the reliability and precision of AWE feedback. Due to AWE's varying error detection rates across English language elements (e.g., grammar, collocations, and mechanics), students selectively evaluated AWE suggestions using their

Table 2. AWE feedback quality (Allen et al., 2018; Foster, 2019; Gao, 2021; Saricaoglu & Bilki, 2021)

AWE Quality Advantages for EFL Classrooms		AWE Quality Limitations for EFL Classrooms	
Key Advantages	Description	Key Limitations	Description
Different types of feedback	AWE provides quality feedback through form-focused, content-based, and integrated feedback	Limited detection of some linguistic errors	AWE systems do not fully detect all linguistic errors, particularly struggling with syntactic improvements and collocation errors.
Supportive role	AWE provides a scaffold that helps students structure high-quality written content and enhances their self-confidence. AWE accelerates the writing process and reduces the number of drafts.	Limited feedback on certain writing aspects	AWE systems may need more feedback on syntactic aspects, which shows that the AWE systems could have limited support for advanced writing.
Accuracy and precision	AWE identifies specific errors, including lexical ones, at a level comparable to professional teachers. This accuracy level enhances the feedback's quality.	Need for Supplementary Feedback Mechanisms	AWE must be supplemented with other feedback mechanisms, such as peer review or teacher feedback, to ensure that all student writing aspects are adequately assessed and improved.
Impact on error reduction	AWE contributes to error reduction in various areas leading to an overall improvement in student writing.	Variable impact on essay improvements	AWE feedback may not always lead to significant improvements in the overall quality of writing. Students might only make minor revisions to their essays.

knowledge of these elements. Their understanding of AWE's strengths and limitations improved with use, aiding their reliability assessment. The study urged informing students about AWE's syntactical and colloquial limitations to boost critical thinking skills, using the system as a cost-effective writing prompt, acknowledging its limitations, and adopting AWE informed by its strengths, drawbacks, student proficiency, pedagogical approaches, and the socio-cultural factors influencing language learning.

Findings by Li et al. (2015) and Bai and Hu (2017) agree regarding the significance of student proficiency and teacher support in AWE adoption in EFL classrooms. The themes explored included instructor views on AWE, their experiences and usage suggestions, feedback approach changes, system evaluation, and student responses. While instructors appreciated AWE's grammar and mechanics support, they found the feedback quality partially satisfying. Students, however, found the corrective feedback beneficial for EFL linguistic accuracy. Therefore, the study underscored the vital role of instructor strategies in implementing AWE and boosting student engagement and awareness of the system's accuracy and reliability. Wang (2021) examined the impact of AWE corrective feedback. Errors within a clause were counted as one, whereas repeated errors across clauses were seen as distinct. The study found that AWE tools failed to detect and correct some logical and syntactic errors. Ranalli and Yamashita (2022) analyzed Grammarly's (<https://app.grammarly.com/>) efficiency in correcting common L2 errors. Grammarly is an AWE tool that uses artificial intelligence and NLP to provide real-time spelling, grammar, and punctuation corrections. It also offers advanced features such as style and tone suggestions. Grammarly flagged over ten times more common L2 problem area errors than MS-NLP. It identified a broad range of errors compared to Microsoft's natural language processing (MS-NLP), including grammatical and syntactical verb forms and inflectional and derivational morphology. The study concluded that Grammarly efficiently addressed complex spelling errors, subject-verb agreement errors, articles, and prepositions, providing timely and effective feedback for L2 students.

Indeed, these studies confirm that AWE showed a high capacity for identifying various errors, including complex spelling, subject-verb agreement, articles, prepositions, and other linguistic errors. Precision rates vary across different error types, like grammar, usage, and mechanics (Bai & Hu, 2017).

In addition, instructional strategies can significantly impact student engagement and ability to evaluate the system's reliability (Li et al., 2015; Bai & Hu, 2017). Further, students assess the reliability of AWE suggestions based on their understanding of grammar, vocabulary, and mechanics. They also make informed decisions about the system's trustworthiness and reliability over usage time (Bai & Hu, 2017). Therefore, AWE's potential limitations should be communicated to students, allowing them to apply critical thinking skills and use other resources (like grammar books and online dictionaries) to verify the system's accuracy and trustworthiness (Bai & Hu, 2017). The system's strengths and drawbacks generally inform the adoption of AWE in EFL classrooms along with the pedagogical approaches shaping language learning activities. Table 3 summarizes the key advantages and limitations of the reliability of AWE feedback discussed in the literature. AWE tools like Grammarly do not have the ability to detect and correct some logical or syntactic errors (Wang, 2021).

AWE Usefulness

Research shows instructor and student confidence in the usefulness of written corrective feedback from AWE systems. AWE is favored for generating reliable scores and offering relevant accuracy feedback (Li et al., 2017; Guo et al., 2021). It also promptly provides consistent feedback on language features upon draft submission (Li et al., 2015), enhancing overall writing quality. AWE's efficacy lies in supplying appropriate language feedback and endless revision and practice opportunities (Guo et

Table 3. AWE feedback reliability (Li et al., 2015; Bai & Hu, 2017; Wang, 2021; Ranalli & Yamashita, 2022)

AWE Quality Advantages for EFL Classrooms		AWE Trust Limitations for EFL Classrooms	
Key Advantages	Description	Key Limitations	Description
Customizable feedback	AWE provides feedback tailored to the specific needs of each student. This feedback can focus on different writing areas, like usage, grammar, and mechanics.	Inconsistent precision and accuracy	There were varying precision and accuracy levels in identifying writing errors across different components of English language writing, such as grammar, collocations, and mechanics. Errors identified by AWE software represent only a small proportion of all errors present in the evaluated texts.
Student trust	Through repeated use, students can gain trust in AWE systems and apply their knowledge to assess the reliability of the feedback. This trust can increase willingness to use the software and follow its suggestions.	Reliance on student knowledge	Students must apply their knowledge of grammar, vocabulary, and mechanics to assess the reliability of AWE feedback. This suggests that students may struggle to use and trust the software without understanding these elements.
Critical thinking development	AWE supports student critical thinking as they use other resources, such as grammar books and online dictionaries, to validate the accuracy and trustworthiness of the system, enhancing their understanding of language and writing mechanics.	Variable quality of feedback	Confidence in AWE cannot be only partial, as the quality of its feedback may not consistently meet user expectations.
Assisting different proficiency levels	AWE supports students with varying proficiency levels, from beginners to advanced learners, tailoring feedback according to their individual needs.	Cultural and social factors	The generalizability and applicability of AWE in different contexts are influenced by the social and cultural factors that shape language learning activities through AWE.
Comparative effectiveness	Many AWE systems have shown advanced capacities in providing effective feedback and detecting various error types.	False positives/negatives	AWE systems might occasionally provide incorrect or misleading feedback. For example, they may flag grammatically correct sentences as incorrect or overlook subtle errors.

al., 2021). It also saves class time, enabling instructors to provide individualized support and content organization feedback.

Using a mixed-method design, Lee (2020) studied Criterion feedback's long-term impact on two Korean undergraduate student English writing proficiency over a year. Criterion (<https://www.ets.org/criterion.html>) is an AWE service developed by the Educational Testing Service (ETS) that uses NLP technology to provide instant essay feedback. It aids in developing writing skills by offering detailed diagnostics on grammar, usage, mechanics, style, organization, and development. The study concluded that AWE improved writing proficiency and enhanced their confidence in their writing skills due to Criterion's perceived usefulness. Students found that it helped improve fluency and grammatical complexity, decreased drafting time, and promoted attention to grammar, form, and usage, aiding in accurate linguistic expression.

Koh (2017) investigated AWE's usefulness at different stages of process-based writing performance. Results indicated that students preferred immediate first draft feedback, with AWE usage decreasing towards the final draft. Students appreciated real-time feedback and content enhancement. AWE allowed more focus on higher-order writing skills, enhancing cognitive capacity for idea elaboration and text readability. It helped resolve language-related issues and provided appropriate grammar feedback. The system also enabled the identification of frequent errors or weaknesses, and students found the feedback improved their English writing ability by providing timely support.

This study's results imply that blending AWE with instructor feedback in process-based writing pedagogy enhances EFL writing performance. AWE's instant feedback on language issues allows students to devote more time to content development, encouraging overall writing improvement by helping them expand and reformulate ideas. However, AWE feedback should be used alongside instructor feedback to ensure students do not overly focus on language-related issues in the early writing stages.

Cheng's (2017) study assessed L2 student writing quality improvements through AWE. The experimental group significantly outperformed the control group, and focus group surveys highlighted the perceived value of AWE in reflective writing. The results suggest that AWE is a helpful tool that can enhance students' reflective L2 learning skills. Parra and Calero (2019) evaluated the effects of implementing free AWE tools on English Teacher Training Programme student writing performance. Ecuadorian undergraduate students used Grammarly to complement teacher feedback. The study reported significant improvements in student writing, comparing the pre-test and post-test scores. The AWE tools effectively enhanced the language learning process, language instruction, and assessment. After receiving direct feedback, learners made fewer errors. The results underscored the positive impact of AWE tools on student writing improvement. In addition to improving writing skills, using AWE tools also boosted student motivation to revise their work and fostered a greater sense of autonomy. Based on these findings, the researchers recommend that teachers give students additional practice opportunities and use the software as a sandbox to test new content composition and text structuring strategies.

These studies provide important insights into the usefulness of AWE systems in improving English writing proficiency, especially in EFL settings. Both instructors and students strongly believe in the usefulness of written corrective feedback provided by AWE systems due to their reliable scoring and relevant corrective feedback (Li et al., 2017; Guo et al., 2021). AWE systems provide immediate and consistent feedback on language-related features after students submit their drafts, positively influencing overall writing quality (Li et al., 2015). AWE systems allow students to revise and practice writing repeatedly and provide crucial language-related feedback. These practice opportunities, in turn, help instructors save class time, which can then be utilized for individualized support and feedback on content and organization (Guo et al., 2021).

AWE has demonstrated long-term benefits in improving writing proficiency, fluency, grammatical complexity, and decreased composing time. It also fosters self-directed learning and improves student confidence in their writing skills (Lee, 2020). Students prefer receiving immediate feedback during

the drafting process and appreciated AWE’s ability to help them resolve language-related issues and focus on higher-order writing skills (Koh, 2017). AWE can enhance students’ reflective L2 learning skills and improve their writing scores (Cheng, 2017). AWE tools significantly improve writing performance by reducing errors and enhancing student motivation to revise their work, fostering a sense of autonomy (Parra & Calero, 2019). AWE showed a high capacity for identifying various errors, including complex spelling, subject-verb agreement, articles, prepositions, and other linguistic errors.

Strategies for Implementing AWE Feedback in EFL Classrooms

The reviewed studies showed that AWE is an effective tool for EFLs to improve their writing skills. It provides immediate feedback on grammar, spelling, and other aspects of written language. AWE feedback can help EFL students identify areas of improvement in their writing and guide them in correcting mistakes. Additionally, automated writing feedback can help EFLs become more aware of the conventions of written English, such as sentence structure and word choice. By providing students with personalized feedback on their writing, automated writing feedback can help them become better writers and more confident communicators in English. AWE can help students to organize their thoughts, structure their arguments, and identify areas of improvement in their writing.

Table 4. Usefulness of AWE Feedback

AWE Usefulness Aspects for EFL Classrooms		AWE Usefulness Limitations for EFL Classrooms	
Key Advantages	Description	Key Limitations	Description
Reliable and valid scoring	AWE system consistently delivers dependable scores, enhancing its credibility. It provides immediate, actionable feedback enhances the speed of the learning process.	Focus on form over content	AWE systems do not support providing meaningful feedback on content, ideas, and argument structure.
Timely feedback	AWE provides immediate feedback on language-related features after students submit writing drafts and aids students in immediate revisions and adjustments.	Not suited for all types of writing	AWE systems are designed for specific types of writing (often academic) and may be less effective or accurate for other types, such as creative writing, poetry, and prose.
Improvement in overall writing quality	AWE positively influences overall writing quality by helping students revise their work based on the feedback provided.	Not adapted to all writing styles	AWE systems are typically designed with a specific “standard” of writing, which may not encompass all writing styles or genres.
Saves instructors time	AWE allows instructors to save time by automating some feedback processes and enables them to utilize this time for one-on-one individualized support and feedback on content and organization.	Potential decrease in teacher–student interaction	AWE could decrease one-on-one feedback sessions between teachers and students, negatively affecting the learning experience.
Unlimited opportunities to practice	Students have unlimited opportunities to revise and practice their writing, aiding continuous improvement and learning. AWE can be a sandbox for students to try new strategies in composing content and structuring text.	Standardization and bias	AWE systems are programmed with specific rules and standards, which can create bias and a lack of adaptation to different writing styles, genres, and cultural expressions.
Enhances proficiency and fluency	AWE system improved the fluency and grammatical complexity in EFL writing, showcasing its usefulness in ESL contexts.	Limited understanding of context	AWE systems need to fully understand the nuances and context of a written piece. They need to be more capable of interpreting a text as a human reader would, which can lead to misunderstandings or inaccurate feedback.

Students should take advantage of the feedback provided by AWE to identify areas of improvement in their writing. Therefore, the most critical synthesized conclusion is that the success of AWE depends highly on the instructor's careful instructional design of the writing activity while implementing AWE (Saricaoglu & Bilki, 2021).

The instructor's role is pivotal in regulating and designing the pre- and post-writing activity tasks to ensure the exploitation of the full potential of the AWE as a pedagogical asset to improve EFL writing. The reviewed studies reflected several insightful recommendations for EFL classroom strategies using AWE that can be categorized through the three phases of writing, including the phase of pre-using AWE in the class, during the AWE activity, and finally, post-using AWE. These strategies can maximize the effectiveness of AWE and improve their academic performance overall.

Pre-AWE Usage

Firstly, it is crucial to establish clear objectives for applying AWE in the classroom. Secondly, selecting the most suitable AWE software program that aligns well with the proficiency level of EFLs is crucial. Moreover, training students to use and leverage AWE systems effectively can significantly mitigate the learning curve associated with these tools and enhance their overall usability (Xu & Zhang, 2021). The studies under review also highlighted the prewriting phase as a critical stage as teachers can bolster students' writing skills by implementing AWE to enhance student self-confidence and readiness to write proficiently in a second language.

Furthermore, combining prewriting strategies and AWE tools effectively enhances student writing abilities. Brainstorming and mind mapping, facilitated by AWE tools, help students organize and clarify their thoughts, contributing to better writing outcomes (Mayer, 2019). Outlining, another writing skill, enables early identification and correction of grammar, style, or structure issues (Xu & Zhang, 2021). Focused practice, encouraged by teachers using AWE tools, provides targeted training on specific prewriting areas that need improvement (Anderson, 2022). Peer review activities and AWE feedback offer a more comprehensive revision process during the drafting phase (Cho & MacArthur, 2011). In addition, using AWE tools on model texts or example essays provides practical learning experiences and helps students experiment with the system and understand the application of feedback in their writing (Stapleton & Wu, 2020). Finally, the gradual introduction of complexity in writing tasks builds student confidence and competence in using AWE tools, thus improving their writing skills (Hartshorn & Evans, 2015).

During the Usage of AWE

During the implementation of AWE, instructors must maintain a balance between content and grammatical correctness when assessing student written assignments while using AWE. Instructors should guide students to concentrate on the details and the overall discourse structure. Engaging cognitively in the reviewing process and navigating through the learning cycle enhances the depth of understanding. Rewriting activities mainly fortify metacognitive reflection of writing skills, improving student performance during the writing activity. In addition, highlighting the negative evidence in students' writing can be a powerful tool to augment engagement with the task and foster the negotiation of meaning. It is also beneficial for instructors to illustrate the practical usage of AWE. By demonstrating how to navigate the software and interpret the provided feedback, instructors can empower students to use these tools more effectively. Moreover, it is recommended that students use AWE in the drafting stage, applying the tool on several drafts before submission. This repeated use ensures comprehensive error correction and increases the likelihood of producing a final paper with high English quality.

In addition, several strategies can effectively incorporate AWE during the writing process to optimize its impact (Li et al., 2015). Encouraging student self-evaluation, for instance, can foster learner independence and facilitate understanding and rectifying their mistakes (Li et al., 2015). Integration of peer review activity after using AWE can provide multiple perspectives on student

writing as peers can highlight content and organization issues, while AWE can identify more technical errors (Cho & MacArthur, 2010). It is also crucial to differentiate and sequence feedback based on each student's proficiency level and learning needs, starting with higher-order errors and progressing to lower-order ones (Hyland & Hyland, 2006; Heift & Schulze, 2007). Furthermore, combining AWE with instructor feedback offers a more comprehensive picture to students, aiding their writing development and enhancing the quality of work produced (Dikli, 2006). If adopted effectively, such strategies can considerably improve the outcome of AWE tools during the writing process.

After Using AWE

Following the implementation of AWE, educators must employ various strategies to ensure the effective utilization and integration of the tool. The support of teacher feedback is essential to complement AWE's input and address the inherent limitations in AWE's cognitive information-processing model. Educators must emphasize that AWE's feedback is designed to foster improvement in student writing abilities and should not be the only method relied upon for skill development (Koltovskaia, 2020). One crucial step is incorporating explicit discussions about AWE feedback into the curriculum to help students understand and act on the feedback (Ware & Warschauer, 2006). This dialogue-driven approach promotes a more profound comprehension of the feedback and its relevance to their writing (Ranalli et al., 2017).

Furthermore, AWE should also be used as a reflective tool, encouraging students to scrutinize the feedback provided and apply it to improve future writing endeavors (Dikli, 2006). It is crucial to provide opportunities for learners to revise their work based on AWE feedback, fostering a deeper understanding of writing mechanics (Hyland & Hyland, 2006). Teachers should consider conducting formative assessments alongside AWE usage to evaluate student progress in real time and identify areas for further reinforcement (Hattie & Timperley, 2007). Implementing peer review sessions after AWE activity can also be an effective strategy, fostering a more collaborative learning environment where students can discuss and interpret AWE feedback together (Liu & Kunnan, 2016).

Moreover, providing adequate practice time with AWE before using it for formal assessment is essential to ensure students can effectively leverage the tool in their learning (Heift & Schulze, 2007). Lastly, while AWE is a valuable teaching aid, educators should use it sparingly to avoid students' over-reliance on technology, which could potentially hamper the long-term development of their writing skills (Li et al., 2015). Maintaining a balance between technology-assisted learning and traditional teaching methods is crucial to optimize learning outcomes.

CONCLUSION

AWE is an effective tool for EFL learners to improve their writing skills; it can help them identify areas of improvement in their writing and provide guidance on correcting mistakes. Additionally, automated writing feedback can help EFL students become more aware of the conventions of written English, such as sentence structure and word choice. By providing students with personalized feedback on their writing, AWE can enhance their self-directed learning, improve their writing skills, and be confident communicators in English. This study explores AWE's effectiveness in improving EFL students writing in English. Further, the study discussed AWE feedback quality, the reliability of AWE, and the usability of these systems supporting EFL writing skills. Finally, this study discussed the most important strategies these research studies recommend when implementing AWE feedback in EFL classrooms. Future research reviews can explore the impact of these strategies and others in optimizing AWE usage and their effectiveness in leveraging the benefits of automated feedback and enhancing feedback quality and usefulness.

The implications and suggestions from the reviewed studies shed light on the strengths and gaps in AWE research. One significant implication is the consensus on AWE's potential to enhance the writing proficiency of EFL students. Its capacity to provide immediate, personalized feedback with

the advantage of reducing teacher workload affirms its value as a powerful tool for EFL instruction (Chen & Cheng, 2008; Barrot, 2021).

However, while the effectiveness and benefits of AWE have been widely explored, research suggests that some areas require further attention. The most critical is the exploration of the pedagogical aspects of AWE feedback. More specifically, maintaining the balance between focusing on the product and the writing process remains a pivotal point in the discussion (Koltovskaia, 2020). Further research is also required to examine how AWE's new features, such as generative writing tools based on artificial intelligence (AI) technologies, can improve the writing process. Additionally, more studies are needed to examine user perceptions and attitudes toward AWE. Understanding student experiences and engagement with the technology will provide invaluable insights into how AWE can be better tailored to meet learner needs. Future studies can also examine how student proficiency levels can allow for a more targeted application of AWE. Lastly, the need for human intervention and the role of teacher feedback and instructional implementation strategies in conjunction with AWE is an important area that needs further exploration.

FUNDING

The authors extend their appreciation to the Deputyship for Research and Innovation, Ministry of Education in Saudi Arabia, for funding this research work through the project number IFPRC-039-126-2020 and King Abdulaziz University, DSR, Jeddah, Saudi Arabia.

CONFLICT OF INTEREST

The authors of this publication declare there is no conflict of interest.

REFERENCES

- Aliakaniotis, D., Yannakoudakis, H., & Rei, M. (2016). Automatic text scoring using neural networks [Preprint 1606.04289]. *arXiv*.
- Allen, L., Likens, A., & McNamara, D. (2018). A multi-dimensional analysis of writing flexibility in an automated writing evaluation system. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, 380–388. doi:10.1145/3170358.3170404
- Anderson, T. (2022). *The theory and practice of online learning*. Athabasca University Press.
- Bai, L., & Hu, G. (2017). In the face of fallible AWE feedback: How do students respond? *Educational Psychology*, 37(1), 67–81. doi:10.1080/01443410.2016.1223275
- Barrot, J. S. (2021). Automated writing evaluation in L2 classrooms: The effect of teacher mediation on student receptivity and text revision. *Computer Assisted Language Learning*, 34(1), 1–24.
- Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: The challenge of design. *Assessment & Evaluation in Higher Education*, 38(6), 698–712. doi:10.1080/02602938.2012.691462
- Burstein, J., Chodorow, M., & Leacock, C. (2004). Automated essay evaluation: The Criterion online writing service. *AI Magazine*, 25, 27–36.
- Chapelle, C. A., Cotos, E., & Lee, J. (2015). Validity arguments for diagnostic assessment using automated writing evaluation. *Language Testing*, 32(3), 385–405. doi:10.1177/0265532214565386
- Chen, C. F. E., & Cheng, W. Y. E. (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.
- Cheng, G. (2017). The impact of online automated feedback on students' reflective journal writing in an EFL course. *The Internet and Higher Education*, 34(1), 18–27. doi:10.1016/j.iheduc.2017.04.002
- Cheng, X., & Zhang, J. (2021). Sustaining university English as a foreign language learners' writing performance through the provision of comprehensive written corrective feedback. *Sustainability (Basel)*, 13(15), 1–19. doi:10.3390/su13158192
- Cho, K., & MacArthur, C. A. (2010). Student revision with peer and expert reviewing. *Learning and Instruction*, 20(4), 328–338. doi:10.1016/j.learninstruc.2009.08.006
- Cho, K., & MacArthur, C. A. (2011). Learning by reviewing. *Journal of Educational Psychology*, 103(1), 73–84. doi:10.1037/a0021950
- Chou, H. C., Moslehpour, M., & Yang, C. (2016). My access and writing error corrections of EFL college preintermediate students. *International Journal of Education*, 8(1), 144–161. doi:10.5296/ije.v8i1.9209
- Dikli, S. (2006). An overview of automated scoring of essays. *The Journal of Technology, Learning, and Assessment*, 5(1).
- El-Ebyary, K., & Windeatt, S. (2010). The impact of computer-based feedback on students' written work. *International Journal of English Studies*, 10(2), 121–142. doi:10.6018/ijes/2010/2/119231
- Foster, S. (2019). What barriers do students perceive to engagement with automated immediate formative feedback? *Journal of Interactive Media in Education*, 1(15), 1–5. doi:10.5334/jime.516
- Gao, J. (2021). Exploring the feedback quality of an automated writing evaluation system Pigai. *International Journal of Emerging Technologies in Learning*, 16(11), 322–330. doi:10.3991/ijet.v16i11.19657
- Grimes, D., & Warschauer, M. (2010). Utility in a fallible tool: A multi-site case study of automated writing evaluation. *The Journal of Technology, Learning, and Assessment*, 8(6), 1–43.
- Guo, Q., Feng, R., & Hua, Y. (2021). How effectively can EFL students use automated written corrective feedback (AWCF) in research writing? *Computer Assisted Language Learning*, 1–20.

- Hartshorn, K. J., Evans, N. W., Merrill, P. F., Sudweeks, R. R., Strong-Krause, D., & Anderson, N. J. (2015). The effects of dynamic corrective feedback on ESL writing accuracy. *TESOL Quarterly*, 49(1), 84–109. doi:10.5054/tq.2010.213781
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. doi:10.3102/003465430298487
- Heift, T., & Schulze, M. (2007). *Errors and intelligence in computer-assisted language learning: Parsers and pedagogues*. Routledge. doi:10.4324/9780203012215
- Hibert, A. I. (2019). Systematic literature review of automated writing evaluation as a formative learning tool. *Proceedings of the European Conference on Technology Enhanced Learning*, 199–212. doi:10.1007/978-3-030-29736-7_15
- Huang, S., & Renandya, W. A. (2020). Exploring the integration of automated feedback among lower-proficiency EFL learners. *Innovation in Language Learning and Teaching*, 14(1), 15–26. doi:10.1080/17501229.2018.1471083
- Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching*, 39(2), 83–101. doi:10.1017/S0261444806003399
- Hyland, K., & Hyland, F. (2019). Contexts and issues in feedback on L2 writing. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 1–22). Cambridge University Press. doi:10.1017/9781108635547.003
- Kerr, P. (2020). *Giving feedback to language learners*. Cambridge University Press.
- Koh, W. (2017). Effective applications of automated writing feedback in process-based writing instruction. *English Teaching*, 72(3), 91–118. doi:10.15858/engtea.72.3.201709.91
- Koltovskaia, T. (2020). The impact of automated feedback on revision behavior in L2 writing. *Journal of Second Language Writing*, 48, 100704.
- Lee, Y. J. (2020). The long-term effect of automated writing evaluation feedback on writing development. *English Teaching*, 75(1), 67–92. doi:10.15858/engtea.75.1.202003.67
- 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. doi:10.1016/j.jslw.2014.10.004
- Li, Z., Feng, H.-H., & Saricaoglu, A. (2017). The short-term and long-term effects of AWE feedback on ESL students' development of grammatical accuracy. *CALICO Journal*, 34(3), 355–375. doi:10.1558/cj.26382
- Li, Z., Link, S., Ma, H., Yang, H., & Hegelheimer, V. (2015). The role of automated writing evaluation holistic scores in the ESL classroom. *System*, 53, 16–27.
- Liu, J., & Kunnan, A. J. (2016). Towards a learning-oriented second language writing assessment: Validity evidence from automated essay scoring practices. *Assessment in Education: Principles, Policy & Practice*, 23(3), 316–334.
- Mayer, R. E. (2019). Multimedia learning. *Psychology of Learning and Motivation*, 41, 85–139. doi:10.1016/S0079-7421(02)80005-6
- Mohsen, M. (2019). The effectiveness of using a hybrid mode of automated writing evaluation system on EFL students' writing. *Teaching English with Technology*, 19, 118–131.
- Parra, G., & Calero, S. (2019). Automated Writing evaluation tools in the improvement of the writing skill. *International Journal of Instruction*, 12(2), 209–226. doi:10.29333/iji.2019.12214a
- Ranalli, J., Link, S., & 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. doi:10.1080/01443410.2015.1136407
- Ranalli, J., & Yamashita, T. (2022). Automated written corrective feedback: Error correction performance and timing of delivery. *Language Learning & Technology*, 26(1), 1–25. <http://hdl.handle.net/10125/73465>

Saricaoglu, A., & Bilki, Z. (2021). Voluntary use of automated writing evaluation by content course students. *ReCALL*, 33(3), 265–s277. doi:10.1017/S0958344021000021

Shermis, G., & Diao, S. (2008). *The impact of automated essay scoring on writing outcomes*. <https://www.semanticscholar.org/paper/The-Impact-of-Automated-Essay-Scoring-on-Writing-Shermis-Garvan/8649af9ac442c62fdf05b0db5935595b887da1e0>

Stapleton, P., & Wu, Y. (2020). Assessing the quality of activities in a self-access language learning center. *Language Teaching Research*, 4(2), 103–123.

Stevenson, M., & Phakiti, A. (2013). The effects of computer-generated feedback on the quality of writing. *Assessing Writing*, 19, 51–65. doi:10.1016/j.asw.2013.11.007

Tang, J., & Rich, C. S. (2017). Automated writing evaluation in an EFL setting: Lessons from China. *The JALT CALL Journal*, 13(2), 117–146. doi:10.29140/jaltcall.v13n2.215

Tang, J., & Wu, Y. (2017). Impact of AWE rubrics and automated assessment on EFL writing instruction. *International Journal of Computer-Assisted Language Learning and Teaching*, 7(2), 58–s74. doi:10.4018/IJCALLT.2017040104

Wang, M., Cheng, C. X., & Chen, S. W. (2006). Contribution of morphological awareness to Chinese–English biliteracy acquisition. *Journal of Educational Psychology*, 98(3), 542–553. doi:10.1037/0022-0663.98.3.542

Wang, S. (2021). Understanding the effects of automated writing corrective feedback on L2 writing accuracy across proficiency levels. *Frontiers in Educational Research*, 4(11), 94–98. doi:10.25236/FER.2021.041117

Ware, P., & Warschauer, M. (2006). Electronic feedback and second language writing. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues*. Cambridge University Press. doi:10.1017/CBO9781139524742.008

Warschauer, M., & Grimes, D. (2008). Automated writing assessment in the classroom. *Pedagogies*, 3(1), 22–36. doi:10.1080/15544800701771580

Wilson, J., & Roscoe, R. D. (2019). Automated writing evaluation and feedback: Multiple metrics of efficacy. *Journal of Educational Computing Research*. Advance online publication. doi:10.1177/0735633119830764

Xu, J., & Zhang, S. (2021). Understanding AWE feedback and English writing of learners with different proficiency levels in an EFL classroom: A sociocultural perspective. *The Asia-Pacific Education Researcher*. Advance online publication. doi:10.1007/s40299-021-00577-7

Yang, H., & Buckendahl, C. W. (2018). Using automated writing evaluation to reduce grammar errors in writing. *Assessing Writing*, 36, 16–27.

Zhai, N., & Ma, X. (2021). Automated writing evaluation (AWE) feedback: A systematic investigation of college students' acceptance. *Computer Assisted Language Learning*, 61, 1–26.

Zhang, S. (2021). Review of automated writing evaluation systems. *Journal of China Computer-Assisted Language Learning*, 1(1), 170–176. doi:10.1515/jccall-2021-2007

Tahani Aldosemani is an associate professor of Educational Technology and Instructional Design at Prince Sattam bin Abdulaziz University, Saudi Arabia. She earned her PhD from the University of Wyoming, USA in 2014. Dr. Aldosemani's research interests include distance online education tools, resources and approaches, instructional design, open educational courses and resources, virtual environments, and skills development.

Hussein Assalahi is an Associate Professor of TESOL at the English Language Institute at King Abdulaziz University. He was awarded a Doctor of Education in TESOL from the University of Exeter, UK in 2016. Dr. Assalahi's particular interests include second language acquisition, computer assisted language learning and language teacher professional development.

Areej Alhothali is an associate professor in the faculty of computer science and information technology at King Abdul-Aziz University. She earned her master's and Ph.D. degrees in computer science (artificial intelligence) from the University of Waterloo, Canada. Her research interest lies in the areas of machine learning, deep learning, natural language processing, intelligent agent systems, affective computing, and sentiment analysis.

Maram Albsisi is a master's student in Computer Science at King Abdulaziz University, Saudi Arabia. Her research interest includes artificial intelligence, data science and data visualizations. Currently, Maram is conducting her thesis on using AI to optimize personalized treatment strategies. She aims to develop AI models to improve healthcare in Saudi Arabia upon graduating with her master's degree.