



AI-Driven Assessment in English as a Foreign Language (EFL) Writing: Balancing Autonomy and Support in the Digital Age

Linda BAGHDADI^{1*} & Dr. Khadidja Samira ZITOUNI²

Type: Full Article. Received: 16/10/2025, Accepted: 18/04/2026

Abstract: Objective: This research aims to explore the perceived effectiveness of AI-powered tools in assessing the writing skills of English as a Foreign Language (EFL) learners. Specifically, it investigates the integration of tools such as Grammarly, QuillBot, and ChatGPT into the writing assessment process. The study addresses the following research questions: (1) To what extent do EFL students and instructors perceive AI tools as effective in assessing writing skills? (2) What are the perceived benefits and challenges? (3) How do participants view the tension between promoting learner autonomy and the risk of overdependence on such tools? **Methodology:** The study targeted 50 university-level Algerian EFL students and instructors, employing an exploratory mixed-methods design and collecting data through an online survey with both closed-ended and open-ended questions. The survey covers areas such as practicality, accuracy, and feedback quality, reflecting the perceived pedagogical value of these tools. **Key Findings:** The findings can be concluded in the following: (1) Both instructors and students hold a general positive perception toward AI assessment. (2) Participants value AI-driven assessment for its immediate personalized feedback, support for learner autonomy and writing improvement, and assistance in reducing workload. (3) Participants are concerned about AI's inaccurate feedback, lack of contextual awareness, and learners' overreliance. (4) Participants perceive AI Assessment as a tool with dual impact, depending on whether it is used appropriately or not, calling for remedial efforts. **Conclusions:** AI is perceived as an effective aid for instructor assessment, offering valuable support for learner autonomy and writing development. However, some concerns are growing about learners' overdependence and faulty feedback. **Recommendations:** A balanced approach through collaborating AI and instructors in writing assessment is recommended. Finally, the study findings contribute by providing valuable insights to the evolving landscape of AI in language education.

Keywords: Algeria, Assessment, AI, EFL, Learner Autonomy, Writing Assessment.

دور الذكاء الاصطناعي في تقييم مهارات الكتابة لدى متعلمي الإنجليزية كلغة أجنبية: نحو توازن بين الدعم واستقلالية المتعلم في العصر الرقمي

ليندة بغداددي^{1*} ود. خديجة سميرة زيتوني²

تاريخ التسليم: 2025/10/16، تاريخ القبول: 2026/04/18

المخلص: الهدف: تسعى هذه الدراسة إلى استكشاف فعالية أدوات الذكاء الاصطناعي في تقييم مهارات الكتابة لدى متعلمي اللغة الإنجليزية كلغة أجنبية، من خلال التركيز على أدوات مثل Grammarly و QuillBot و ChatGPT. وتتناول الدراسة الأسئلة التالية: 1 كيف ينظر الطلاب والمعلمون إلى فعالية أدوات الذكاء الاصطناعي في تقييم مهارات الكتابة؟ 2 ما هي الفوائد والتحديات التي يواجهونها عند الاستخدام؟ 3 كيف يوازن المعلمون والطلاب بين تعزيز استقلالية المتعلم وخطر الاعتماد المفرط؟ **المنهجية:** استهدفت الدراسة 50 طالب ومعلم إنجليزية في المرحلة الجامعية، واعتمدت على تصميم استكشافي باستخدام منهجية تجمع بين الكمي والنوعي. تم جمع البيانات من خلال استبيان إلكتروني تضمن أسئلة مغلقة وأخرى مفتوحة، بهدف الحصول على صورة شاملة حول فعالية هذه الأدوات، ودقتها، وجودة تقييماتها، وأثرها التربوي. **أهم النتائج:** أظهرت النتائج ما يلي: 1 معظم الطلاب والمعلمين عبّروا عن نظرة إيجابية تجاه استخدام أدوات الذكاء الاصطناعي في تقييم الكتابة. 2 تتمثل أبرز مزايا هذه الأدوات في قدرتها على تقديم تقييمات فورية ومخصصة، ومساعدتها في تحسين مهارات الكتابة، ودعم استقلالية المتعلم، إلى جانب تخفيف العبء عن المعلمين. 3 في المقابل، أبدى المشاركون مخاوف تتعلق بعدم دقة بعض التقييمات، وغياب الفهم السياقي، إضافة إلى خطر اعتماد الطلاب الزائد على هذه الأدوات. 4 تشير النتائج إلى أن استخدام الذكاء الاصطناعي في التقييم يمكن أن يكون ذا أثر مزدوج بحسب كيفية استخدامه، مما يستدعي توجيهها واضحا وممارسة رشيدة. **الاستنتاجات:** يُعد الذكاء الاصطناعي أداة مفيدة تدعم المعلمين في تقييم الكتابة من خلال تقديم ملاحظات فورية ومساعدة الطلاب على تطوير مهاراتهم. ومع ذلك، هناك حاجة إلى توعية الطلاب والمعلمين بمخاطر الاعتماد الكلي على هذه الأدوات، خصوصا في ظل بعض أوجه القصور المتعلقة بدقتها وفهمها للسياق. **التوصيات:** توصي الدراسة بدمج دور المعلم مع الذكاء الاصطناعي لتحقيق توازن فعال في عملية التقييم. وتساهم هذه النتائج في تعزيز فهمنا للدور المتنامي للذكاء الاصطناعي في تعليم اللغات وتقييمها.

الكلمات المفتاحية: الجزائر، التقييم، الذكاء الاصطناعي، اللغة الإنجليزية كلغة أجنبية، استقلالية المتعلم، تقييم الكتابة.

1 Laboratory of SEPRADIS, Department of English Language and Literature, Faculty of Letters & Foreign Languages, University Centre of Barika, Amdoukal Road, Barika, 05001, Algeria.

* Corresponding author: linda.baghdadi@cu-barika.dz

ORCID: <https://orcid.org/0009-0005-0104-4464>

2 Laboratory of Sodidlem, Department of English Language and Literature, Faculty of Letters & Foreign Languages, University Centre of Barika, Amdoukal Road, Barika, 05001, Algeria.

khadidja.zitouni@cu-barika.dz

ORCID: <https://orcid.org/0000-0002-3773-5110>

1 مخبر SEPRADIS، قسم اللغة الإنجليزية وأدبها، كلية الآداب واللغات الأجنبية، المركز الجامعي لبريكة، طريق أموكال، بريكة، 05001، الجزائر.

* الباحث المراسل: linda.baghdadi@cu-barika.dz

<https://orcid.org/0009-0005-0104-4464>

2 مخبر Sodidlem، قسم اللغة الإنجليزية وأدبها، كلية الآداب واللغات الأجنبية، المركز الجامعي لبريكة، طريق أموكال، بريكة، 05001، الجزائر.

khadidja.zitouni@cu-barika.dz

<https://orcid.org/0000-0002-3773-5110>

Background

Assessment is a multifaceted process that requires intensive efforts to ensure accurate measurement and comprehensive analysis. In the context of language learning and teaching, the inherently complex nature of language acquisition makes assessing learner progress even more complicated. This challenge calls for innovative strategies to facilitate the assessment process, including the integration of Artificial Intelligence (AI) (Richardson and Clesham, 2021). Popenici and Kerr (2017) defined AI as computing systems that can learn, adapt, synthesize, and solve problems in a human-like manner. The goal of these systems is to carry out responsibilities in the same way humans do (Abdul Rahman et al., 2023), serving as a viable alternative that enhances time efficiency and provides personalized feedback.

Moreover, writing is such a difficult skill to develop as it requires considerable cognitive and language proficiency (Faller, 2018; Maharani, 2018). According to Boukhelkhal (2025), AI tools are constantly being developed to meet the established principles of language learning, teaching, and assessment. In the context of assessment, AI plays a significant role by providing formative feedback (Abdul Rahman et al., 2023) that enables progress tracking. Stošić and Malyuga (2024) propose that AI has facilitated the creation of tools that provide personalized assessments tailored to learners' proficiency and needs, allowing for immediate error detection and correction (Malyuga et al., 2016). The advantages of using AI for assessment are not limited to test personalization and time efficiency, as it can also enhance both objectivity and accuracy in analysis. Assessors can minimize personal bias and ensure fairness in assessment by relying on standardized criteria and data-driven evaluation methods, ensuring that all

students are measured on the basis of consistent benchmarks.

However, some challenges are faced when incorporating AI in assessment, including the need for extensive data, struggles in understanding contexts, and ethical and privacy concerns (Stošić and Malyuga, 2024). In Algerian English as a Foreign Language (EFL) classrooms, the pedagogical effectiveness of AI-driven assessment has yet to be fully explored. While numerous studies have investigated the integration of AI in education, limited data uncover the perceptions of both students and instructors towards using AI technologies such as Grammarly, QuillBot, and ChatGPT for assessment. Exploring this matter can offer a comprehensive understanding of the benefits and challenges of AI-driven assessment and provide further insights into the conflict between learner autonomy and over-reliance.

For that, the current research aims to explore the perceived effectiveness of AI-driven assessment in Algerian EFL classrooms. Precisely, it aims to collect insights from both EFL instructors and students toward the benefits and challenges of using AI technologies for assessment. Additionally, it aims to explore the conflict between AI's role in promoting learner autonomy and the risk of overreliance. To achieve that, it aims to answer the following questions:

Q1: To what extent do EFL students and instructors perceive AI tools as effective in assessing writing skills?

Q2: What are the perceived benefits and challenges of using AI tools for writing assessment?

Q3: How do participants view the tension between promoting learner autonomy and the risk of overdependence on such tools?

The study holds significance by providing valuable insights into how AI-driven assessment could be implemented effectively, reducing workload and time pressure, and enhancing assessment accuracy. Its findings add to the body of knowledge, inform future research, and offer implications for teaching practices and policymakers. Overall, the study underscores the pedagogical potential of AI-driven assessment in promoting personalized learning experiences and the faced difficulties.

2. Literature review

People may think of AI as a recent concept, but it goes as far back as 1956, with numerous studies investigating the ability of computers to simulate human intelligence. Artificial Intelligence progressed at an accelerated pace that enabled its impact to reach and reform almost all disciplines, including education. The latter underwent a transformation that reshaped its teaching and learning approaches. In the context of writing assessment, AI contributes significantly by providing formative feedback and assessment (Abdul Rahman et al., 2023). Popenici and Kerr (2017) defined AI as systems that are capable of imitating human intelligence, including learning, adapting, synthesizing, and problem-solving. These skills are translated as personalized learning paths in educational contexts. AI-driven assessment tools provide students with tailored input based on their individual needs and proficiency level (Stošić and Malyuga, 2024). This alliance between AI features and pedagogical needs promoted the implementation of AI for assessment across various EFL classrooms.

Moreover, applications such as Grammarly, QuillBot, and ChatGPT are widely used for writing in academic contexts. The tools offer various features, including grammar correction, paraphrasing, and feedback, supporting self-assessment. Natural Language

Processing (NLP) algorithms, which are used by the mentioned tools, allow for the analysis of complex language structures and a personalized approach in evaluation (Stošić and Malyuga, 2024). This feature of AI, in addition to other features such as machine translations, adaptive learning, and speech recognition, allows for feedback tailored to the individual weaknesses and strengths (Bakhodirov and Rahmanova, 2023, as cited in Stošić and Malyuga, 2024). According to Gorozhanov and Guseynova (2020), AI is able to distinguish contextual variations beyond just grammatical accuracy. For that, AI can design tests and exercises for specific needs and focus on particular grammar aspects that require further improvement (Stošić and Malyuga, 2024). Ultimately, AI has the potential to promote formative assessment and provide personalized feedback.

Furthermore, learner autonomy has arisen as an area of interest when discussing the integration of AI into the process of language learning and teaching. According to Benson (2007), learner autonomy is “learners’ ability to take charge of their own learning”, and according to Kohonen et al. (2001), its promotion includes making decisions about material selection, reflection, and identifying strengths and weaknesses. AI in this context serves as a beneficial tool that can assist learners in taking control over their learning process and being less dependent on their instructors. According to Patty (2023), AI enables progress tracking, collaboration, and personalized learning paths. Moreover, Wei (2023) proposes that integrating AI into language instruction could promote students’ motivation and autonomy and improve language learning outcomes. With the appropriate use, AI enables learners to make choices and decisions, enhance their competencies, and develop confidence to engage in real-world situations.

Despite the numerous benefits of AI, students' misuse of Artificial Intelligence may take the wrong turn and hinder their progress. In 2023, UNESCO conducted a study entitled *Guidance for Generative AI in Education and Research*, which highlights some concerns about human intellectual growth in relation to the use of Generative AI in education. It proposes that "Gen AI systems in education may limit learners' autonomy and agency by providing predetermined solutions or narrowing the range of possible learning experiences" (p.36). Overreliance on AI costs learners opportunities where they could have learned from their mistakes and engaged in real-world contexts with authentic language. Spector and Ma (2019) propose that over-dependence on AI may negatively impact learners' critical thinking abilities. Additionally, Kerma's (2025) study findings suggest that excessive use of AI tools may impair students' engagement and critical thinking skills. This contrast between AI assistance in promoting learner autonomy and learners' overreliance on it calls for further research.

This dual impact of AI has arisen as an area of interest in recent research. Hamamra et al (2025a) examined how Palestinian university instructors perceive the integration of AI into educational settings, revealing both enthusiasm and concerns. While their findings proposed potential personalization and enhanced academic achievement, they also highlighted the risk of learning depersonalization caused by over-reliance on AI and negligence of teacher-student interactions. The latter issue calls for a human-centered integration of AI into education where interpersonal interactions are prioritized, and critical thinking, emotional intelligence, and ethical reasoning are nurtured (Hamamra et al., 2025b). In fact, the advantages of AI cannot be denied; however,

problems appear when AI is used as a substitute rather than an assistant. The latter causes many harms, including weakened academic achievement, motivational decline, ethical risks, creativity suppression, and much more (Alhur et al., 2025). For that, it is recommended that institutions provide clear guidelines, ongoing practical training, and strong support systems that would ensure proper and responsible use of AI (Khlaif et al., 2025).

In the Algerian educational context, several previous studies investigated the integration of AI and its prospects. Guemide and Sahraoui (2023) explored Algerian EFL students' perceptions, which highlighted AI's immediate feedback and ability to foster engagement, leading to better learning outcomes. Additionally, Boudouaia et al. (2024) conducted an experimental study investigating the usefulness of ChatGPT-4 in students' EFL writing. The findings revealed a positive impact on writing improvement and perceived usefulness from the participants. On the other side, Benaicha and Semmoud (2024) and Merdassi and Belmekki (2024) examined the attitudes of EFL teachers from various Algerian universities, and the findings revealed concerns about students' over-reliance on AI, which may gradually weaken their critical thinking. Finally, Boukhelkhal (2025) investigated students' perceptions and practices related to AI-based learning. The results indicated that students hold positive views on AI-assisted tools (primarily ChatGPT) due to their immediate feedback, language enhancement, and personalized learning experiences. However, concerns remain about faulty content, over-reliance, and decreased critical thinking skills.

While several studies were conducted on AI in relation to EFL writing in Algerian classrooms, attention was directed largely toward students' attitudes and the perceived

benefits and challenges. However, the aspect of assessment, the formative one, particularly, has yet to be thoroughly explored. Also, there is a lack of studies examining both students' and instructors' perceptions, which is necessary to ensure more comprehensive insights. In addition, limited research has explored the conflict between learner autonomy and potential overreliance on AI tools in the context of assessment. The current study aims to fill the gap by exploring the perceived effectiveness of AI-driven assessment in Algerian EFL classrooms, providing nuanced insights into AI's position in language learning and teaching. It also aims to collect data from the dual perspectives of instructors and students to gain a holistic understanding, critically examining how a balance between independence and AI assistance could be navigated. Additionally, the Technology Acceptance Model (TAM) by Fred D. Davis (1989) will be used as the theoretical lens guiding the interpretation of findings, particularly its core constructs of perceived usefulness and perceived ease of use, which offer a structured framework for understanding the degree to which participants believe AI assists in enhancing their performance and the degree to which it requires little effort. Finally, the study findings offer implications for classroom practices, policymakers, and future research.

Method

Study Design: The current study employed an exploratory research design with a mixed-methods approach suitable for investigating a relatively underexplored topic, which is AI-driven assessment in the Algerian EFL context.

Study population and Setting: The study targeted Algerian EFL instructors and post-graduate students from multiple universities to obtain insights from both sides. The

universities were as follows: University Centre of Barika, University of Batna 2, National Higher School of Bouzaréah (Algiers), University of Mohamed Khider (Biskra). Additionally, data were collected through an online survey from 7 July 2025 to 27 July 2025.

Study Sample: A sample of 50 participants was selected through a snowball technique. The latter was selected due to the exploratory nature of the study, which prioritizes depth of perception over statistical generalizability. The inclusion criteria for selecting participants were mainly prior familiarity with and use of at least one AI writing tool, ensuring that responses reflected valid perceptions that are grounded in concrete experience rather than misperceptions or prejudice. The target population was narrowed to two groups known for relatively higher engagement with digital pedagogies and emerging technologies due to their recent and ongoing exposure: EFL post-graduate students (Master's and PhD level) and novice university instructors with one to three years of teaching experience. Recruitment began with the researchers' professional networks, and participants were then invited to refer colleagues meeting the same inclusion criteria. The survey was designed using Google Forms and distributed electronically via Email, a formal platform commonly used for academic communication. It should be noted that recruiting beyond the achieved sample proved challenging, as low response rates are a common limitation in voluntary academic surveys. By the end, a sample of 25 EFL post-graduate students and 25 novice university instructors was composed, with an equal distribution across both groups intended to allow for a balanced comparison of their perceptions.

Background variables such as age or gender were not included as the study is grounded in perceptions shaped by AI use rather than

demographics. And since the selection of participants was not limited to one particular institution, the participants belonged to multiple universities across Algeria.

Data Collection Tool: The data were collected through an online questionnaire. The tool contained sixteen questions divided into six sections, covering participants' profiles, familiarity and awareness of usage, perceived effectiveness with both quantitative and qualitative insights, and recommendations. The tool administered contained both closed-ended and open-ended questions to collect mixed data.

Validity and Reliability of the Tool: The questionnaire was self-developed by the researchers and informed by the existing body of literature on AI, writing assessment, formative feedback, and learner autonomy (content validity). To further ensure the validity of the research, several fellow researchers were asked to review the data gathering tool and provide their feedback (face validity). The remarks provided were taken into account, and adjustments were made to enhance the tool's efficiency. Additionally, pilot testing was conducted to enhance the clarity of questionnaire items.

The questionnaire consisted of five distinct sections serving different measurement purposes. Conducting a reliability analysis was not applicable or appropriate for the first two sections due to the nature of their items, which were designed to describe participants' backgrounds and behavioral patterns. To explain further, the first section captures demographic and background information, indicating participant role and frequency of AI tool usage with single categorical and ordinal variables. Additionally, the second section comprised checklist items allowing participants to select the AI tools they used and their purposes of use, reflecting

independent behavioral choices rather than a one unified dimensional scale. The third section contained six Likert-scale attitudinal items measuring participants' perceptions of AI writing tools. In this case, reliability analysis was conducted and yielded a Cronbach's alpha of .85, indicating strong internal consistency. Item-total correlations ranged from .582 to .717, and alpha-if-item-deleted values ranged from .821 to .850, confirming that all six items contributed meaningfully to the scale and that no item deletion would improve reliability. The fourth and fifth sections consisted of open-ended questions that were designed to capture qualitative insights grounded in personal experiences, which by nature do not lend themselves to reliability analysis in the psychometric sense.

Data Analysis: Quantitative data collected through closed-ended questions underwent descriptive statistical analysis via SPSS and were presented in terms of frequencies, percentages, means, modes, medians, standard deviations, and interquartile ranges. Additionally, the Chi-square test of independence was used to examine associations between group (instructors vs. students) and binary variables (e.g., tool usage), with Cramér's V as a measure of effect size. For ordinal Likert-scale items, the non-parametric Mann–Whitney U test was applied to compare distributions between the two independent groups.

Furthermore, thematic analysis was employed to examine qualitative data extracted from open-ended questions, following an open and axial coding procedure. In the first phase, all responses to the open-ended questions were read repeatedly, and initial codes were examined inductively. In the second phase, related open codes were grouped into broader categories, which were refined into core themes. A codebook was

developed inductively from the data and definitions, and representative quotations were used to evidence each theme. To enhance the trustworthiness of the coding process, which included one coder, a transparent coding procedure, iterative theme refinement, and peer debriefing, was implemented. The latter maintained inter-subjectivity through having the second author review the coded data and provide critical feedback and notes, ensuring consistency in interpretation.

Results

This section aims to present, analyze, and interpret the data collected from fifty university-level Algerian EFL instructors and students. The instrument used for data collection was an online survey with both open-ended and closed-ended questions, obtaining both quantitative and qualitative insights. This section is guided by the survey's items, undergoing a descriptive statistical analysis or a thematic one, depending on the type of question.

Item 1: What is your role?

Table (1). Participants' Role (n = 50).

Category	N	%
Instructor	25	50%
Student	25	50%

After a short introduction, the survey started exploring participants' profiles. It started with a descriptive overview, which aimed to identify the role of each participant, as represented in Table (1). As fifty participants took part in the study, half of them (50%) were instructors, while the other half (50%) were students. This balance in distribution ensures equal representation of both teachers' and students' perspectives.

Item 2: How often do you use AI writing tools such as Grammarly, QuillBot, and ChatGPT?

Table (2). Usage Frequency (n = 50).

Category	N	%
Never	2	4%
Rarely	10	20%
Weekly	26	52%
Daily	12	24%

The survey continued to build participants' profiles, particularly on how often they use AI writing tools. As represented in Table (2), participants were asked to select one of the provided usage frequencies. Half of the participants (52%) claimed to use AI writing tools weekly, while 12 of them (24%) claimed daily usage. Also, a considerable number of 10 participants (20%) rarely use AI for writing assistance, and only 2 do not use it at all (4%). This diversity in engagement levels enables obtaining non-biased insights from both frequent users and those with limited experience. Additionally, the Chi-square test of independence indicates that there is no significant association in the usage patterns among instructors and students, as the value was $\chi^2 (3, N = 50) = 4.55, p = .208 > .05$. The latter indicates a similar pattern of usage between both sides. However, two cells (25% of cells) had expected counts less than 5 (minimum = 1.0), violating the chi-square assumption. To account for this, a Monte Carlo simulation with 10,000 sampled tables was conducted, confirming no significant association ($p = 1.000$). The effect size, Cramér's $V = 0.084$, indicated a very small association between group and response selection. Generally, the results indicate a weekly usage (mode=3, median=3, mean=2.96); also, the STD of .78 suggests that the responses do not vary much among all participants.

Item 3: Which of the following Tools have you used for writing or assessment? (Select all that apply)

Table (3). Tools (n = 50).

Tool	Yes, I use it		No, I do not	
	N	%	N	%
Grammarly	20	40%	30	60%
QuillBot	13	26%	37	74%
ChatGPT	44	88%	6	12%
Other	9	18%	41	82%

Table (3) above represents participants' familiarity with different AI tools in the context of assessment. The third item was a multiple-choice question where participants could select more than one tool. To ensure a subtle analysis, each tool was coded into a separate binary variable (Yes/No). The most commonly used tool was ChatGPT with 44 participants (88%), followed by Grammarly with 20 participants (40%). Moreover, 13 participants (26%) used QuillBot for writing assessment, while 9 participants (18%) claimed to use other AI tools, including DeepSeek, Gemini, and Perplexity. The findings indicate the popularity of ChatGPT among participants, although Grammarly is older. And even though there are various other AI tools, participants' low usage of them indicates limited awareness or accessibility.

Furthermore, Chi-square tests or Fisher's Exact Tests were used to examine whether tool usage differed significantly between instructors and students. Where expected cell counts fell below 5, Fisher's Exact Test was applied in place of the chi-square test. Results indicated no statistically significant association between group and usage of any of the four tools. For Grammarly, $\chi^2(1, N = 50) = 1.33$, $p = .248$, Cramér's $V = .16$; for QuillBot, $\chi^2(1, N = 50) = 0.10$, $p = .747$, Cramér's $V = .05$; for ChatGPT, Fisher's Exact Test $p = .667$, Cramér's $V = .12$; and for other tools, Fisher's Exact Test $p = .138$, Cramér's $V = .26$. Effect sizes were uniformly small across all comparisons, suggesting negligible differences in tool usage patterns between the two groups.

The slightly higher Cramér's V for other tools (.26) tentatively suggests that students may be more open to exploring alternative AI tools than instructors. However, the results indicate no statistically significant difference in general, highlighting similar preferences among students and instructors.

Item 4: What do you primarily use these tools for? (Select all that apply)

Table (4). Purposes (n = 50).

Purpose	Yes, I use it		No, I do not	
	N	%	N	%
Paraphrasing	29	58%	21	42%
Grammar Correction	25	50%	25	50%
Vocabulary Enhancement	21	42%	29	58%
Essay Structure	11	22%	39	78%
Feedback	29	58%	21	42%
Scoring/ Assessment	18	36%	32	64%
Another	3	6%	47	94%

Table (4) represents different aspects of assessment; in other words, particular purposes. The item aimed to investigate the key merits of AI-driven assessment and its usage context. Since participants could select more than one purpose, each purpose was coded into a binary variable (yes/no) to ensure nuanced examination. The most common purposes for using AI tools in writing assessment are paraphrasing and feedback (58%), followed by grammar correction (50%) and vocabulary enhancement (42%). Moreover, 36% of participants use AI for scoring and assessment, with 22% using it for essay structure. Meanwhile, only three participants (6%) claimed to use AI for other purposes, such as generating content.

Chi-square tests were conducted to examine whether the purposes of AI tool use differed

significantly between instructors and students. All expected cell counts exceeded 5 across all purpose categories, satisfying the chi-square assumption, with the exception of the 'other purposes' category, where Fisher's Exact Test was applied. Results indicated no statistically significant association between group and any of the seven purposes examined. None of the seven purpose categories showed a statistically significant association with group (all $p > .05$; Cramér's V ranging from .04 to .12). Additionally, effect sizes were negligible to small across all comparisons (ranging from Cramér's $V = .04$ to .12), indicating that instructors and students did not differ meaningfully in their purposes for using AI writing tools. Overall, the values indicate that both instructors and students share similar purposes when using AI-driven assessment.

In addition, an independent-samples Mann–Whitney U test was conducted to examine the differences among students' and instructors' perceptions across Likert scale items. The results were as follows: *impact on writing accuracy* ($p = .322$), *trust in AI-driven assessment* ($p = .216$), *quality of feedback* ($p = .844$), *academic writing standards* ($p = .164$), *AI feedback vs instructor feedback* ($p = .826$), and *self-assessment* ($p = .358$). The results indicate that there are no statistically significant differences between the two groups across all items, suggesting that students and instructors share similar perceptions concerning the effectiveness of AI-driven assessment.

Item 5: AI tools help improve my / my students' writing accuracy

Table (5). Impact on Writing Accuracy (n = 50).

Category	N	%
Strongly Disagree	1	2%
Disagree	0	0%
Neutral	18	36%
Agree	19	38%
Strongly Agree	12	24%
Median	4	
Interquartile Range	1	

After building participants' profiles and exploring their familiarity and usage context with several AI assessment tools, the survey moved into Likert scales to evaluate the perceived effectiveness. Table (5) represents AI's impact on writing accuracy. The majority of participants hold a positive attitude, with 38% agreeing and 24% strongly agreeing that AI helps improve writing accuracy, 36% remaining neutral, and only one participant (2%) disagreeing. The Median of 4 indicates a positive evaluation. Meanwhile, the IQR of 1 indicates strong agreement among participants. Ultimately, AI tools help improve writing accuracy according to the majority of participants (62%).

Item 6: I trust AI tools for assessing writing tasks

Table (6). Trust in AI-Driven Assessment (n = 50).

Category	N	%
Strongly Disagree	2	4%
Disagree	6	12%
Neutral	17	34%
Agree	17	34%
Strongly Agree	8	16%
Median	3.5	
Interquartile Range	1	

Table (6) represents participants' trust in AI tools for assessing writing tasks. 34% of participants preferred to remain neutral. Another 34% of them agreed, and 16% strongly agreed with the provided statement. However, 12% disagreed, and 4% strongly disagreed. Additionally, the Median is 3.5 with an IQR of 1, highlighting a mixture of uncertainty, trust, and caution. The results indicate a generally moderate level of trust in AI tools when used for writing assessment (50%). But it appears that many participants are skeptical about AI's assessment and whether it is fully accurate and unbiased.

Item 7: AI tools provide timely and useful feedback

Table (7). Quality of Feedback (n = 50).

Category	N	%
Strongly Disagree	1	2%
Disagree	1	2%
Neutral	9	18%
Agree	19	38%
Strongly Agree	20	40%
Median	4	
Interquartile Range	1	

Table (7) summarizes participants' perceptions of AI's feedback. The results indicate positive perceptions, as 38% agreed and 40% strongly agreed that the feedback is timely and useful. Meanwhile, only 18% remained neutral, and 4% disagreed with the statement. The Median is 4 while the IQR is 1, indicating a strong positive evaluation of AI's feedback with strong agreement in answers. To sum up, AI's feedback is timely and useful according to the majority of participants (78%).

Item 8: AI writing feedback aligns with academic writing standards

Table (8). Academic Writing Standards (n = 50).

Category	N	%
Strongly Disagree	1	2%
Disagree	3	6%
Neutral	13	26%
Agree	22	44%
Strongly Agree	11	22%
Median	4	
Interquartile Range	1	

Table (8) shows whether participants believe that AI writing feedback aligns with academic writing standards. The results were as follows: 44% agreed, 22% strongly agreed, 26% neutral, 6% disagreed, and 2% strongly disagreed. As for the Median, it is 4 with an IQR of 1, indicating strong agreement on positive evaluation. Overall, AI's feedback aligns with academic writing standards according to the majority of participants (66%).

Item 9: AI helps identify areas for improvement better than peer/instructor feedback

Table (9). AI feedback vs peer/instructor feedback (n = 50).

Category	N	%
Strongly Disagree	4	8%
Disagree	12	24%
Neutral	13	26%
Agree	15	30%
Strongly Agree	6	12%
Median	3	
Interquartile Range	2	

Table (9) represents whether participants agree that AI helps identify areas for improvement better than peer/instructor

feedback. The results were as follows: 30% of participants agreed, 12% strongly agreed, 26% were neutral, 24% disagreed, and 8% strongly disagreed with the statement. The results indicate mixed perceptions of AI's feedback assistance in improvement. Some view it as an effective alternative to traditional feedback, while others remain uncertain or disagree. The Median of 3 indicates uncertainty, while the IQR of 2 suggests a relatively wide spread of opinions. To sum up, AI's feedback is not superior to peer/instructor feedback according to 58% of the participants.

Item 10: AI tools have improved my ability / my students' ability to self-edit

Table (10). Self-assessment (n = 50).

Category	N	%
Strongly Disagree	3	6%
Disagree	5	10%
Neutral	14	28%
Agree	18	36%
Strongly Agree	10	20%
Median	4	
Interquartile Range	1	

Table 10 represents whether participants agree that AI tools have improved their own or their students' ability to self-edit. The results were as follows: 36% agree, 20% strongly agree, 28% neutral, 10% disagree, and 6% strongly disagree. In addition, the Median of 4 and the IQR of 1 indicate positive perceptions and strong agreement. Ultimately, AI is perceived as helpful in self-assessment according to the majority of participants (56%).

Item 11: In your experience, what are the main benefits of using AI tools in writing assessment?

Since the study adopted a mixed-method approach, the survey encompassed open-ended questions to obtain qualitative data. Question

number eleven aimed to explore the perceived benefits of using AI tools in writing assessment. Participants' answers were summarized into themes, with the most dominant one being *Time Efficiency and Accessibility*. The participants consistently praised AI in relation to time and accessibility. For example, one participant stated that: "AI tools provide immediate/instant feedback". Another participant believed that: "AI tools save a lot of time", and another one additionally emphasized how: "They're accessible anytime". The second theme extracted is *Writing Enhancement*, with many participants emphasizing how AI helps in improving writing quality. In this context, one participant mentioned how: "AI tools refine grammar, punctuation, and paraphrasing", in addition to another one emphasizing how they: "Assist in enhancing vocabulary". Additionally, another participant mentioned how: "They provide powerful wording", indicating that AI is perceived as a reliable writing assistant. The third theme examined from the responses is *Personalization and Objectivity*. Multiple participants highlighted how AI tailors feedback to learners' individual needs, without subjectivity. Participants declared statements such as "Personalized feedback", "Precise feedback based on your needs", "Objective grading by removing human bias", which indicate trust in AI's assessment as an accurate and non-biased one. As for the last theme, it is *Promoted Learner Autonomy and Self-Assessment*. Many responses suggest that AI helps learners take ownership of their learning journey and become independent. The theme is laid in answers such as: "It helps you assess yourself", "Students become aware of their own mistakes", "Learning by doing", and "Self-learning".

Item 12: What challenges or limitations have you encountered when using AI tools?

This item aimed to explore the challenges and limitations encountered when using AI tools. The first theme examined is *Faulty Feedback*, as many participants expressed their concerns regarding AI's inaccurate, biased, and misleading feedback. One participant, for example, stated that: "AI tools sometimes make stuff up", and another one highlighted how "They are not always accurate". Additionally, other participants emphasized how "Some AI tools change the core of the ideas" and "Provide fake information", addressing the need to double-check the information provided. The second theme examined is *Limited and Costly Access*. Participants expressed their frustration with how premium tools have a daily usage limit or require payment/subscription, limiting their usability. Additionally, participants warned from the *Learners' Overreliance* on AI, which would kill creativity and hinder critical thinking development. Lastly, "Lack of Human Insight" was another area concerning the participants. According to them: "AI's feedback lacks human ability", "Misunderstands instructions", and "Not suitable for all contexts", leading to miscommunication and inaccurate assessment.

Item 13: How do you feel about using AI for formative assessment compared to human feedback?

The survey continued to understand how participants perceive AI-driven assessment, particularly its role in formative assessment, compared to human feedback. Four main themes were examined, starting with *AI's Time Efficiency and Accessibility*. Many participants valued AI and stated that it: "Facilitates the process", "Provides instant feedback", and "Saves a lot of time". Participants believed that AI reduces teachers' workload and enhances learners' independence. As for the second theme, it is *AI's Feedback is Complementary to human*

feedback. Many participants perceived AI as a supportive tool rather than a replacement for human feedback and suggested some ways in which they can work together. For example, one participant suggested that: "AI tools should be used to fine-tune students' production, nothing more". Additionally, multiple participants believed that *AI's Feedback Lacks the Depth and Empathy That Exists in Human Feedback*. The latter is valued for its "Empathy that machines can't have" and its "contextual understanding". Last but not least, participants expressed their concerns about *AI's Limitations and Learners' Overreliance on it*. They believe that, regardless of AI's personalized feedback, it may lack context awareness and could be too "General" sometimes. Also, participants expressed their concerns, including how "AI uses highly advanced vocabulary" and how students may over-depend on it in doing their tasks.

Item 14: Do you think AI promotes learner autonomy or causes overdependence? Why?

Additionally, the survey explored AI's educational impact on learner autonomy. The majority of participants believed that AI has a *Dual Potential* and could promote autonomy or cause overreliance, depending on learners' usage. They emphasized how the aftermath is based on: "whether students use AI to learn or as a shortcut", and "whether they use it properly or not". Moreover, several participants believed that *AI Mostly Causes Overreliance and Kills Critical Thinking*. They worry that "Learners may not bother themselves to think" and that "Their critical thinking mind zone starts to be narrowed". Meanwhile, other participants viewed *AI as an Assistant in Promoting Learner Autonomy*. They believed that "if it is used properly, it will enable students to track their progress and improve their writing skills". Finally, the *Need*

for *Rising Students' Awareness and Guiding them on the Appropriate Use of AI* was highlighted by participants as a necessary procedure to ensure proper implementation of AI-driven assessment.

Item 15: How do you think AI and instructors can complement each other in the writing assessment process?

When asked how AI and instructors can complement each other in the writing assessment process, participants suggested *Task Division*. The majority of participants proposed that AI assessment tools can handle surface-level tasks. For example, one participant suggested that: "AI tools can handle grammar, spelling, structure, and plagiarism checking, while instructors can focus on deeper elements such as creativity, context, critical thinking, and argumentation". Another participant highlighted how "combining the two types will create a more efficient and supportive assessment process". Additionally, a large number of participants emphasized *AI as an Assistant, not a replacement for teachers*. They suggested that AI can assess repetitive tasks and save the instructor's time. So, the latter can dedicate more time to providing in-depth and personalized feedback. One of the participants recommended "Asking students to work without AI first, then use it to finalize their work and submit both versions". This way, instructors can assess progress and learning difficulties, and learners' metacognition and autonomy would be promoted.

Item 16: Do you have any recommendations for using AI in writing assessment more effectively?

The responses to Item 16 reveal a set of pedagogical, methodological, and ethical considerations for the integration of AI into writing assessment. By analyzing the

responses, the following themes were identified:

Theme 1: Pedagogical Role of AI

- Participants emphasized that AI assessment tools should assist and support the teaching–learning process rather than replace human judgment or generate entire outputs.
- AI should be employed purposively, with its use adapted to the specific learning and assessment context.

Theme 2: Critical Engagement with AI Feedback

- Respondents noted that AI feedback must be questioned and critically analyzed by both students and instructors.
- They also highlighted the importance of raising awareness about the risks of overreliance on AI, warning that excessive dependence may undermine students' autonomy and writing development.

Theme 3: Training and Guidance

- The need for professional training for instructors and orientation for students was strongly emphasized to ensure effective and appropriate integration of AI tools.
- Teachers were seen as playing a guiding role in monitoring and directing how students engage with AI within writing assessment.

Theme 4: Ethical and Responsible Use

- Participants recommended that AI should be integrated in ways that promote ethical, intellectual, and responsible use, reinforcing academic integrity and critical literacy.

In order to reach deeper insights, the qualitative findings were interpreted through the Technology Acceptance Model (TAM; Davis, 1989). To start with, the participants highlighted the *perceived usefulness of AI*,

addressing its automated/immediate feedback, personalized output, time efficiency, and support for learner autonomy. However, challenges such as faulty feedback, limited access, and learner over-reliance reflect limitations in the *perceived ease of use*. The findings indicate that an effective implication of AI requires a thorough understanding of its capabilities and limitations. Additionally, participants emphasized AI's role as an assistant that can manage repetitive/routine tasks; meanwhile, teachers remain responsible for higher-order, creative assessment. For effective use of AI in writing assessment, formal training, guidance, and critical engagement are recommended.

Discussion

Findings indicate that both students and teachers reported comparable usage patterns and perceptions, suggesting a shared understanding of the role of AI in writing assessment. Additionally, the two sides share similar usage patterns (weekly), similar preferences in AI tools (mostly ChatGPT), and similar purposes. Participants use AI-driven assessment mostly for feedback, paraphrasing, grammar correction, and vocabulary enhancement. In addition, they believe that AI enhances writing accuracy and provides immediate, useful feedback that aligns with academic standards. However, they remain skeptical about its potentially inaccurate, biased assessment. The findings align with those of Boudouaia et al. (2024) and Hamamra et al. (2025b), where ChatGPT is perceived as useful in writing improvement but not a replacement for instructor/peer feedback. AI assessment tools are better suited for proofreading and repetitive tasks.

Findings also indicate that AI holds several perceived benefits when integrated into writing assessment. According to the participants, AI's time efficiency and

accessibility can streamline the assessment process and writing development, aligning with Guemide and Sahraoui's (2023) study findings. Additionally, AI can enhance writing quality through proofreading and provide immediate personalized feedback tailored to learners' needs without human subjectivity. The examined perceived benefits align with those of Boukhelkhal (2025). As reported, AI can support self-assessment and promote both metacognition and autonomy. It can also reduce workload and assist instructors in the assessment process by handling surface-level, repetitive tasks. Nevertheless, Participants expressed concerns about the reliability of AI-generated feedback and the risk of learner overdependence, which affects their intellectual and writing development, as suggested by Boukhelkhal (2025) and Alhur et al. (2025). The feedback provided by AI lacks human empathy and contextual awareness, which makes it too general sometimes. Also, its vocabulary is too advanced for the learner's level sometimes, causing misunderstandings. Ultimately, the findings indicate that AI should be perceived as an assistant, not a replacement for teachers, as its feedback alone is not sufficient.

The study findings align with those of Hamamra et al. (2025a) as they indicate how AI holds a dual potential depending on how it is used. On one hand, AI will assist in enhancing learners' autonomy if they use its feedback to learn, reflect, and find competency gaps. Learners should take advantage of AI's automated feedback to streamline the assessment process and support their writing development. However, they must also engage critically with this feedback, as it may be misleading or inaccurate sometimes. AI-driven assessment encompasses facilities that align with constructivist and student-centered learning approaches. On the other hand, AI can hinder

learners' progress if they misuse it or excessively depend on its generative feedback. Some students may use AI tools as shortcuts to finish entire assignments, rather than employing them as supportive aids in the writing process. Some do not even know how to use these tools appropriately and ask the right prompts, leading to misleading feedback. Moreover, overreliance on AI tools kills creativity, limits writing development, and reduces critical thinking. The concerns examined align with those of Benaicha and Semmoud (2024) and Merdassi and Belmekki (2024).

Finally, the procedures recommended for an effective integration of AI-driven assessment in Algerian EFL classrooms align with those recommended by Khlaif et al. (2025). The participants suggested the following:

- AI assessment tools should exist to assist instructors, not replace them.
- Users should learn to analyze AI's feedback critically and not take it for granted.
- Institutions should raise awareness of the consequences of overreliance on AI that may undermine users' autonomy, critical thinking, and writing development.
- Both instructors and students should be provided with formal training on the appropriate use of AI assessment tools.

Ultimately, AI-driven assessment is an effective learning assistant for writing improvement if incorporated appropriately with human instruction.

Conclusion

This study has shown that Algerian EFL instructors and students perceive AI-driven assessment as an effective learning assistant

for writing improvement if incorporated appropriately with human instruction. AI-driven assessment is valued by participants for its automated personalized feedback, support for learner autonomy and writing improvement, and assistance in reducing teachers' workload. However, some concerns have emerged about AI's inaccurate feedback, lack of contextual awareness, and learners' overreliance. This dual impact is dependent on whether AI assessment tools are used appropriately or not. Therefore, participants emphasized that AI assessment tools must serve as a complementary assistant, not a replacement for human feedback. It is also noted that the lack of awareness of academic integrity risks plagiarism and ethical violations, so institutional policies should promote responsible use of AI. Finally, professional training is recommended by participants for both instructors and students, emphasizing critical thinking and independence.

Some limitations should be acknowledged, including the limited generalizability of findings due to the small sample size and the use of only one instrument for data collection. As for the delimitations, the study explored AI-driven assessment in EFL writing, excluding other language skills, and focused solely on its implications in higher education. Additionally, from the variety of AI tools available for assessment, only three were selected (Grammarly, QuillBot, and ChatGPT) for the study.

The study recommendations mirror those outlined in the Discussion, emphasizing institutional awareness, formal training, and critical engagement with AI feedback. Additionally, further research could conduct a comparative study between traditional assessment and AI-driven assessment.

Investigating the long-term impact of AI's automated feedback on writing development is also possible. And finally, a study exploring the best strategies for integrating AI tools into traditional assessment methods would bring profound understanding into leveraging AI in education.

Disclosure Statement

The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

- **Ethical approval and consent to participate:** Ethical principles were carefully established throughout the research to protect its integrity and the participants' rights. Participation was entirely voluntary, and informed consent was obtained. All participants were informed about the aim of the study before taking the survey and were assured that they could withdraw at any point if they wanted to. Also, their autonomy was respected by making sure that their responses were not influenced by the researcher. Furthermore, confidentiality and anonymity were guaranteed by avoiding the collection of any personal identifiers and using the data collected solely for academic purposes.
- **Availability of data and materials:** The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.
- **Author contribution:** All authors listed have contributed to the work and approved it for publication. The authors have worked together and in an organized manner to accomplish the research. The authors read and approved the final manuscript.
- **Conflict of interest:** The authors declare that they have no competing interests.

- **Funding:** The researchers received no funding for this study. Additionally, this study is an original work that has not been extracted from a master's or doctoral thesis, nor has it been published elsewhere.
- **Acknowledgements:** We want to thank all the participants who provided such valuable insights to our study.

List of Abbreviations

AI: Artificial Intelligence

EFL: English as a Foreign Language

NLP: Natural Language Processing

Open Access

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution, and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third-party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/deed.en>

References

- Abdul Rahman, N. A. B. H., Zulkornain, L. H., Che Mat, A., & Kustati, M. (2023). Assessing Writing Abilities Using AI-Powered Writing Evaluations. *Journal of ASIAN Behavioural Studies*, 8(24). <https://doi.org/10.21834/jabs.v8i24.420>

- Alhur, A. A., Khlaif, Z. N., Hamamra, B., & Hussein, E. (2025). Paradox of AI in higher education: Qualitative inquiry into AI dependency among educators in Palestine. *JMIR Medical Education*. <https://doi.org/10.2196/74947>
- Bakhodirov, O., & Rahmanova, G. (2023, October). The Role of AI in Language Learning Apps. In *International Conference on Higher Education Teaching* (Vol. 1, No. 11, pp. 9–12).
- Benaicha, B., & Semmoud, A. (2024). Investigating Algerian EFL Teachers' Attitudes towards AI Utilization in Language Education. *ATRAS Journal*, 5(3), 130–150. Retrieved from <https://asjp.cerist.dz/en/article/253642>
- Benson, P. (2007). Autonomy in Language Teaching and Learning. *Language Teaching*, 40(1), 21–40. <https://doi.org/10.1017/s0261444806003958>
- Boudouaia, A., Mouas, S., & Kouider, B. (2024). A Study on ChatGPT-4 as an Innovative Approach to Enhancing English as a Foreign Language Writing Learning. *Journal of Educational Computing Research*, 62(6), 1509–1537. <https://doi.org/10.1177/07356331241247465>
- Boukhelkhal, O. (2025). Algerian EFL Students' Perspectives and Practices on AI-Enabled Learning at the University of Medea. *Journal of Studies in Language, Culture, and Society*, 8(1), 93–110. <https://asjp.cerist.dz/en/article/266935>
- Faller, J. M. V. (2018). Grammarly Investigation into EFL Writing Issues Involving Omani Learners. *International Journal of Language & Linguistics*, 5(3). <https://doi.org/10.30845/ijll.v5n316>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- Gorozhanov, A. I., & Guseynova, I. A. (2020). Programming for Specific Purposes in Linguistics: A New Challenge for the Humanitarian Curricula. *Training, Language and Culture*, 4(4), 23–38. <https://doi.org/10.22363/2521-442X-2020-4-4-23-38>
- Guemide, B., & Sahraoui, A. A. (2023). Implementing AI Applications to Improve English Language Learning Among EFL Students in Algeria: A Survey Study from University Students' Perspectives. *International Journal of Literacy and Education*, 3(2), 38–46. Retrieved from <https://www.educationjournal.info/archives/2023.v3.i2.A.146>
- Hamamra, B., Khlaif, Z. N., & Mayaleh, A. (2025a). Between promise and precarity: Palestinian university educators' experiences with AI integration in higher education. *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2025.2547264>
- Hamamra, B., Khlaif, Z. N., Mayaleh, A., & Abu Baker, A. (2025b). A phenomenological examination of educators' experiences with AI integration in Palestinian higher education. *Cogent Education*, 12(1). <https://doi.org/10.1080/2331186X.2025.2526435>
- Kerma, M. (2025). The impact of artificial intelligence on Algerian learners' critical thinking. *ATRAS*, 6(1), 125–136. <https://asjp.cerist.dz/en/article/262181>
- Khlaif, Z. N., Hamamra, B., Bensalem, E., Mitwally, M. A. A., & Sanmugam, M. (2025). Factors influencing educators' technostress while using generative AI: A

- qualitative study. *Technology, Knowledge, and Learning*.
<https://doi.org/10.1007/s10758-025-09885-z>
- Kohonen, V., Jaatinen, R., Kaikkonen, P., & Lehtovaara, J. (2001). *Experiential Learning in Foreign Language Education* (1st ed.). Routledge.
<https://doi.org/10.4324/9781315840505>
- Maharani, M. M. (2018). Graphic Organisers to Improve Students' Writing on Recount Paragraphs. *Metathesis: Journal of English Language, Literature, and Teaching*, 2(2), 211–221.
<https://doi.org/10.31002/metathesis.v2i2.942>
- Malyuga, E. N., Shvets, A., & Tikhomirov, I. (2016). Computer-Based Analysis of Business Communication Language. In *Proceedings of 2016 SAI Computing Conference SAI 2016* (pp. 229–232). IEEE.
<https://doi.org/10.1109/SAI.2016.7555987>
- Merdassi, N., & Belmekki, A. (2024). Striking the Balance of AI Use in EFL Education: Maximizing Benefits and Minimizing Risks. *ATRAS Journal*, 5(3), 69–79.
<https://asjp.cerist.dz/en/article/253638>
- Patty, J. (2023). The Use of AI in Language Learning: What You Need To Know. *Journal Review Pendidikan Dan Pengajaran*, 7(1).
<https://journal.universitaspahlawan.ac.id/index.php/jrpp/article/view/24609>
- Popenici, S. A. D., & Kerr, S. (2017). Exploring the Impact of Artificial Intelligence on Teaching and Learning in Higher Education. *Research and Practice in Technology Enhanced Learning*, 12(1), 22. <https://doi.org/10.1186/s41039-017-0062-8>
- Richardson, M., & Clesham, R. (2021). Rise of the machines? The evolving role of Artificial Intelligence (AI) technologies in high-stakes assessment. *London Review of Education*, 19(1), 1–13.
<https://doi.org/10.14324/LRE.19.1.10>
- Spector, J. M., & Ma, S. (2019). Inquiry and Critical Thinking Skills for The Next Generation: From Artificial Intelligence Back to Human Intelligence. *Smart Learning Environments*, 6(1), 8.
<https://doi.org/10.1186/s40561-019-0088-z>
- Stošić, L., & Malyuga, E. (2024). Application of Artificial Intelligence in Language Skills Testing. *ANGLISTICUM Journal of the Association-Institute for English Language and American Studies*, 13(1), 22–34.
<https://doi.org/10.58885/ijllis.v13i1.22ls>
- UNESCO. (2023). Guidance for Generative AI in Education and Research. *United Nations Educational, Scientific, and Cultural Organization*.
<https://www.oph.fi/sites/default/files/documents/Guidance%20for%20generative%20AI%20in%20education%20and%20research.pdf>
- Wei, L. (2023). Artificial Intelligence in Language Instruction: Impact on English Learning Achievement, L2 Motivation, and Self-Regulated Learning. *Frontiers in Psychology*, 14.
<https://doi.org/10.3389/fpsyg.2023.1261955>

Appendix

Survey

AI-Driven Assessment in English as a Foreign Language (EFL) Writing: Balancing Autonomy and Support in the Digital Age

Introduction

You are kindly invited to participate in a research survey that aims to explore how university-level EFL students and instructors perceive the use of AI tools (Grammarly, QuillBot, ChatGPT) in assessing writing skills. The survey will take approximately 10-15 minutes to complete, answering a combination of multiple-choice and open-ended questions.

Your responses are anonymous, and all data will be used for academic purposes only. Participation is voluntary, and you may exit the survey at any time. Your participation is truly appreciated. Thank you.

Section One: Participants' Profile

1. What is your role?

Instructor ☐

Student ☐

2. How often do you use AI writing tools such as Grammarly, QuillBot, and ChatGPT?

Daily ☐

Weekly ☐

Rarely ☐

Never ☐

Section Two: Participants' Familiarity and Awareness of Usage

3. Which of the following tools have you used for writing or assessment? (Select all that apply)

Grammarly ☐

QuillBot ☐

ChatGPT ☐

Other:

4. What do you primarily use these tools for? (Select all that apply)

Paraphrasing ☐

Grammar correction ☐

Vocabulary enhancement ☐

Essay structure ☐

Feedback ☐

Scoring/Assessment ☐

Another:.....

Section Three: Quantitative Insights

Please select from the scale what best reflects your opinion.

5. AI tools help improve my / my students' writing accuracy.

Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree ☐

6. I trust AI tools for assessing writing tasks.

Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree ☐

7. AI tools provide timely and useful feedback.

Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree ☐

8. AI writing feedback aligns with academic writing standards.

Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree ☐

9. AI helps identify areas for improvement better than peer/instructor feedback.

Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree ☐

10. AI tools have improved my ability / my students' ability to self-edit.

Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree ☐

Section Four: Qualitative Insights

11. In your experience, what are the main benefits of using AI tools in writing assessment?

-
12. What challenges or limitations have you encountered when using AI tools?
-
13. How do you feel about using AI for formative assessment compared to human feedback?
-
14. Do you think AI promotes learner autonomy or causes overdependence? Why?

-
15. How do you think AI and instructors can complement each other in the writing assessment process?
-
- Section Five: Suggestions for Improvement**
16. Do you have any recommendations for using AI in writing assessment more effectively?
-