AI-ENABLED FORMATIVE EVALUATION IN EFL WRITING: POSSIBILITIES AND TEACHING CHALLENGES

Balancing Automation with Human Judgement for Continuous Learning

Rafida Nur Azizah, Rudi Hartono, Sri Wahyuni
Master's Program in English Language Education
Semarang State University
Semarang, Indonesia
rafida3796@students.unnes.ac.id

Abstract

Given this increasing trend of integrating artificial intelligence (AI) into education, the methods used for assessment in EFL writing classrooms are also changing. This paper explores the uses and challenges of integrating AI in formative assessment of writing. Drawing upon recent qualitative research in formative assessment and AI-enabled feedback, this review is grounded on a conceptual framework through literature published during the period from 2020 to 2025. This paper focuses on the various ways in which AI tools can be used to create writing as well as how this may impact assessment practices.

The analysis' results reveal that AI tools can make formative assessments more efficient and accessible. They allow students to receive quicker feedback, edit more frequently, and be a more active part of their learning journey. In many cases, AI also supports educators in providing them with automatic but uniform feedback on common writing problems that arise in large classes. However, they still have certain limitations with more complex writing-related aspects like argumentation, audience-context consideration, and idea expression. In addition, there are fears that students may become too reliant on AI or use it without a sufficient understanding of the feedback provided.

Thus, this work promotes an equilibrated model where AI is used to provide basic feedback, but human educators carry on the responsibility of nurturing knowledge comprehension and interpretation. To finish, the use of AI in EFL writing assessment is beneficial, as long as it needs to be wisely adopted so that it serves meaningful learning instead of replacement.

Keywords: artificial intelligence, formative assessment, EFL writing, automated feedback, AI-assisted learning

Introduction

Assessment of EFL writing has, for a long time, been associated with numerical grades, correction marks and the final written product. Feedback is usually given after a task and for the sake of assessment rather than learning in many classrooms. These practices often limit the opportunities for students to re-revise their texts in meaningful and active engagement with feedback. According to Nurhayati (2020), there is a more considerable option for formative assessment since it provides feedback that influences both teaching and learning so students are active, reflective, and engaged in the reformulation of their writing. Likewise, Wijaya (2022) also states that formative assessment improves better quality and motivation for writing image by encouraging a more comprehensive process of writing.

The theoretical perspective of this study is based primarily on formative assessment theory, where assessment is seen as an ongoing activity designed to promote learning through feedback and opportunities for revisions (Black & Wiliam, 2009). It also builds on self-regulated learning, where learners ask for feedback and improve their own performance through reflection. These theoretical perspectives are relevant because they have the potential to support iterative feedback cycles and promote learner autonomy, which can be beneficial

for writing development through AI-powered tools.

The importance of formative assessment becomes even more visible when writing is understood as a process rather than a single product. EFL students usually need support not only after they finish writing, but also while they are planning, drafting, revising, and editing their texts. In this sense, formative assessment helps students notice their weaknesses, revise their work, and gradually develop control over language use. The same principle is reflected in recent discussions of AI-assisted writing, where digital tools are not only used to correct errors but also to support students during the writing process. For instance, AI writing tools can assist students in generating ideas, revising sentences, enhancing organization, and obtaining immediate suggestions prior to receiving final feedback from the teacher. This relevance of AI to formative assessment lies in its ability to facilitate more frequent feedback opportunities during the writing process, rather than confining assessment to a final evaluation.

The rise of artificial intelligence has introduced new possibilities into this area. AI-powered tools can provide instant feedback, identify linguistic issues, support drafting and revision processes, and reduce teachers' workload in writing assessment. These developments are particularly relevant in EFL classrooms, where large student numbers and limited instructional time often make individualized writing feedback difficult to sustain. Huang and Winkel (n.d.) explain that AI-powered automated feedback tools can handle lower-order concerns such as grammar, references, and language conventions, allowing teachers to focus more on higher-order issues such as reasoning and content. In a similar vein, Nazli et al. (2025) report that AI-powered corrective feedback can improve students' writing performance by offering immediate and personalized correction, although important limitations remain.

Several recent studies show that AI has become increasingly important in language learning and writing pedagogy. Hussain et al. (2024) explain that AI-powered personalized learning can identify learner needs, provide customized instruction, and support more adaptive language learning experiences. Aysu (2024) also notes that AI tools in language learning offer benefits such as instant feedback, accessibility, motivation, and personalized learning, although they may also create problems related to ethical use, limited human interaction, and critical thinking. In the Indonesian context, Hasibuan et al. (2023) argue that AI applications have transformed language learning by supporting personalized learning, assessment, and student-centered practices. These studies suggest that AI is no longer only an optional technological addition; it is becoming part of the broader transformation of language education.

These developments suggest that AI has considerable potential to support formative assessment in EFL writing. Unlike summative assessment, formative assessment is development-oriented: it seeks to provide feedback that helps learners improve during the writing process, not merely after it. This orientation aligns well with the strengths of many AI tools, which can provide rapid feedback across multiple drafts and support iterative revision. Williyan et al. (2025) found that AI-powered pedagogy in tertiary ELT supported real-time feedback, reflective revision, and argument development in argumentative essay writing. Alnemrat et al. (2025) also show that AI-generated feedback can improve EFL argumentative writing performance comparably to teacher-generated feedback, suggesting that AI may become a scalable complement in writing classrooms.

At the same time, the integration of AI into formative assessment is not without challenges. AI feedback is often strongest in handling lower-order language issues, but weaker in addressing higher-order concerns such as argument quality, reasoning, and rhetorical effectiveness. Tahir et al. (2025) found that AI-powered writing assistants can support writing development, but their effect on critical thinking remains uncertain, especially because lower-proficiency learners may depend on AI for surface-level revisions rather than deeper engagement with ideas.

However, AI-powered formative assessment should not be seen only from the perspective of technological efficiency. Writing assessment is a pedagogical activity that involves interpretation, judgment, and understanding of learner development. Although AI can provide fast and repeated feedback, teachers still need to decide whether the feedback is appropriate for the student's level, writing purpose, and classroom objectives. This is important because writing quality involves not only grammar and mechanics, but also content, organization, coherence, argumentation, voice, and rhetorical appropriateness. Consequently, the primary concern is not whether AI can evaluate writing, but rather how AI can be incorporated into formative assessment without undermining teacher judgment and learner reflection.

Despite the increasing volume of research on AI in language learning and EFL writing, there has been limited focus on how AI can be systematically integrated into formative assessment frameworks. Most research tends to emphasize AI as a writing tool instead of a pedagogically sound assessment strategy. Thus, a more precise conceptualization of AI-enabled formative assessment remains necessary.

This raises important pedagogical questions: How should AI-powered tools be positioned in formative assessment? What opportunities do they offer for EFL writing development? What challenges do they create for teachers and learners?

Based on these concerns, this paper explores AI-powered formative assessment in EFL writing with a focus on both its opportunities and its pedagogical challenges. It argues that AI can strengthen formative assessment when used as an assistive feedback tool, but that its effectiveness depends on careful pedagogical integration and continued teacher oversight.

This paper contributes to the existing literature by synthesizing research on AI in EFL writing and formative assessment, and by proposing a hybrid model that integrates AI-powered feedback with teacher-mediated assessment practices.

Methodology

This study employs a qualitative conceptual paper design based on library research. A conceptual paper does not aim to generate empirical data, but to develop a theoretical and analytical discussion by synthesizing existing literature (X). This approach is appropriate because the study aims to examine how AI-powered tools can be positioned within formative assessment in EFL writing.

The literature reviewed in this paper consists of recent journal articles focusing on AI in language learning, AI-powered feedback, and formative assessment in EFL writing. The sources were selected based on their relevance to the topic, recency (2020–2025), and contribution to discussions on AI-assisted writing and assessment. These studies were identified through academic databases such as Google Scholar and other indexing platforms.

The literature selected for this paper falls into three main areas. First, the paper draws on studies of formative assessment in EFL writing, especially those emphasizing improvement-oriented assessment, learner

reflection, revision, and the pedagogical value of feedback. Nurhayati (2020) and Wijaya (2022) are particularly central in this area because they frame formative assessment as a process that supports writing development, learner engagement, and more meaningful teaching practices. These studies are used to establish the formative assessment foundation of the paper before the discussion moves to AI-supported writing feedback.

Second, the paper reviews studies on AI-powered writing feedback, AI writing tools, and EFL writing development. This group includes the sources already used in the paper, such as Nazli et al. (2025), Huang and Winkel (n.d.), Tahir et al. (2025), and Alnemrat et al. (2025), because they discuss AI corrective feedback, automated feedback systems, AI writing assistants, and comparisons between AI-generated and teacher-generated feedback. In addition, this paper also draws on studies from the fifteen reviewed articles, including Harunasari (2023), Al Mahmud (2023), Wale and Kassahun (2024), Song and Song (2023), Marzuki et al. (2023), and Alzahrani and Alotaibi (2024). These studies are relevant because they examine specific AI writing tools such as ChatGPT, Wordtune, Writerly, Google Docs, and other AI writing applications in EFL writing contexts.

Third, the paper includes literature on AI in language learning, AI-supported assessment, and AI-based language evaluation. Williyan et al. (2025) and Valerio (2025) are used because they connect AI tools with broader EFL writing pedagogy, especially in argumentative and academic writing contexts. The paper also includes broader studies on AI-powered language learning and personalized instruction, such as Hussain et al. (2024), Aysu (2024), and Hasibuan et al. (2023). In addition, studies by Parsa et al. (2021), Sheikh (2024), Mousa (2025), Li et al. (2024), Manggiasih et al. (2023), and Williams (2025) are included because they discuss AI-powered assessment, automated feedback, AI-assisted marking, formative assessment, language proficiency evaluation, and the limitations of AI-based language tools.

These three groups of literature were selected because they collectively support the focus of this paper. The first group provides the theoretical foundation of formative assessment in EFL writing. The second group connects AI directly to writing feedback, writing tools, and EFL writing development. The third group expands the discussion to AI-supported assessment, language learning, and the limitations of automated evaluation. Together, the earlier sources and the fifteen additional articles provide a basis for discussing both the opportunities and pedagogical challenges of AI-powered formative assessment in EFL writing.

The analysis proceeded in three stages. First, the literature was read to identify how formative assessment is conceptualized in EFL writing. Second, the literature on AI-powered tools was examined to identify their main functions, affordances, and limitations in writing assessment. Third, these two strands were synthesized to develop an argument about the opportunities and pedagogical challenges of AI-powered formative assessment in EFL writing.

Because this is a conceptual paper, the purpose of the analysis is not to generalize statistically but to develop a coherent academic argument supported by existing literature.

The synthesis was conducted thematically. First, the studies were grouped according to their main focus: AI in language learning, AI in EFL writing, AI feedback and assessment, and AI-based language evaluation. Second, the findings of each study were compared to identify recurring themes, such as instant feedback, revision support, learner autonomy, teacher workload, feedback accuracy, ethical

concerns, and the need for teacher mediation. Third, these themes were connected to the concept of formative assessment in EFL writing. This thematic synthesis is suitable for a conceptual paper because the aim is not to measure the effectiveness of one specific AI tool, but to explain how different studies collectively inform the discussion of AI-powered formative assessment.

Findings and Discussion

1. AI-powered tools expand the possibilities of formative assessment

One of the clearest opportunities offered by AI-powered tools in EFL writing is the expansion of formative assessment practices. Formative assessment depends on timely, usable feedback that helps students revise and improve their work while learning is still in progress. In conventional writing classrooms, however, this ideal is often difficult to achieve because teachers face time constraints, heavy workloads, and large class sizes. Nurhayati (2020) notes that formative assessment is valuable because it supports improvement and allows students to reflect on their writing, but such practices require sustained teacher input. AI tools help address this challenge by making feedback faster and more scalable.

The expansion of formative assessment through AI can also be connected to personalized learning. Hussain et al. (2024) argue that AI-powered personalized learning can assess proficiency levels, identify learning gaps, and deliver instructional content according to individual learner needs. In EFL writing, this means that AI feedback can potentially respond to different student problems in different ways. A student who struggles with grammar may receive corrective feedback on sentence structure, while another student who struggles with vocabulary may receive suggestions for more precise word choices. This kind of individualized feedback is difficult for teachers to provide consistently in large classrooms. Therefore, AI can support formative assessment by making individualized feedback more available and more frequent.

Table 1. Opportunities and Challenges of Integrating AI in Formative Assessment

Category	*Key Findings and Descriptions*	*Relevant Sources (Selected)*
Opportunities	**Efficiency & Scalability:** Offers instant feedback in large classes where teacher time is limited.	Nurhayati (2020); Huang & Winkel (n.d.)
	Learner Autonomy: Fosters self-regulated learning through iterative drafting and revision.	Williyan et al. (2025); Nazli et al. (2025)
	Motivation: Increases student engagement with feedback while ideas are still fresh.	Song & Song (2023)
Challenges	**Overreliance:** Risk of students accepting AI feedback mechanically without reflection.	Harunasari (2023); Tahir et al. (2025)
	Cognitive Limitations: Difficulty in evaluating complex logic,	Marzuki et al. (2023); Williams (2025)

audience nuance, and voice.	
Ethical Issues: Concerns over data privacy, academic integrity, and AI literacy.	Sheikh (2024); Mousa (2025)

Similarly, Nazli et al. (2025) found that AI-powered corrective feedback can improve writing performance by offering immediate, tailored, and stable feedback, which in turn supports learner engagement and self-regulated learning. This is particularly relevant for formative assessment because delayed feedback often reduces students' willingness or ability to revise meaningfully. When students receive feedback instantly, they can revise while their ideas are still fresh, making the assessment process more dynamic and responsive.

Huang and Winkel (n.d.) similarly emphasize that automated feedback tools can provide instant formative feedback on lower-order writing concerns and fit within existing learning management systems. This kind of integration strengthens formative assessment by making feedback more continuous and embedded in the writing process rather than leaving it as an isolated post-task event. In this way, AI supports one of the central goals of formative assessment: helping learners improve through iterative feedback cycles.

2. AI supports revision, learner autonomy, and self-regulated writing

Another significant opportunity exists in the capacity of AI-powered tools to support revision and foster learner autonomy. Effective writing development relies on revision; however, many students find it challenging to revise meaningfully unless they receive feedback that is specific, accessible, and actionable. AI tools can help here by giving students repeated feedback opportunities across drafting stages.

Williyan et al. (2025) found that AI-powered pedagogy in argumentative essay writing enhanced real-time feedback, peer review, reflective revision, and the structuring of better arguments. This suggests that AI tools can strengthen the recursive nature of writing by supporting revision as an ongoing process. Valerio (2025) likewise found that AI-powered tools, particularly when combined with traditional instruction, improved writing competency and long-term retention of argumentative writing skills.

This finding is also supported by Song and Song (2023), who found that ChatGPT-assisted language learning improved EFL students' academic writing skills and motivation. Their study showed improvement in several aspects of writing, including organization, coherence, grammar, and vocabulary. This is relevant to formative assessment because motivation and writing improvement are closely connected. When students receive immediate and accessible feedback, they may become more inclined to revise their drafts and engage with the writing process. Consequently, ChatGPT can be perceived not only as a writing assistance tool but also as a formative resource that motivates students to enhance their writing through continuous feedback and revision.

This point is also supported by studies on specific AI writing tools. Al Mahmud (2023) found that Wordtune helped EFL students improve their writing at lexical and syntactic levels by offering rewritten versions of students' sentences. This is important for formative assessment because students can compare their original sentences with AI-generated alternatives and notice how word choice, sentence structure, and clarity can be improved. Wale and Kassahun (2024) also found that the integration of Writerly and Google Docs improved EFL writing instruction because

students were able to receive support, revise their writing, and collaborate more effectively. These findings show that AI tools can make revision more visible and more manageable for learners.

These findings align closely with formative assessment principles. Formative assessment works best when students are not passive recipients of scores but active participants in monitoring and improving their own learning. Nazli et al. (2025) explicitly connect AI feedback with self-regulated learning, suggesting that AI can promote learner responsibility in revision. In this sense, AI-powered formative assessment may help students become more independent writers who can notice problems, respond to suggestions, and improve their texts through multiple drafting cycles.

However, learner autonomy in this context should not be understood in a simplistic manner. AI can encourage independent revision practices, but only if students are able to interpret and use feedback critically. Otherwise, AI may lead to passive acceptance rather than reflective learning.

For this reason, learner autonomy should be connected to AI literacy. Aysu (2024) explains that students need to understand how AI tools work, how to access them, how to write effective prompts, how to check AI output, and how to incorporate AI-generated suggestions responsibly. In EFL writing assessment, this means that students should not only be allowed to use AI, but should also be trained to evaluate AI feedback. They need to learn that AI suggestions are not always correct, complete, or suitable for the intended writing purpose. Thus, AI-supported formative assessment should encompass reflection activities, such as prompting students to clarify which AI recommendations they accepted, rejected, or altered.

3. AI reduces teacher workload and improves assessment scalability

A persistent challenge in EFL writing classrooms is the difficulty of providing individualized feedback to many students. Writing assessment is labor-intensive, and the ideal of formative assessment can become difficult to maintain in large classes. One important advantage of AI-powered tools is therefore their potential to improve assessment scalability.

The issue of teacher workload is also discussed in studies on AI-powered assessment. Sheikh (2024) explains that AI assessment systems can automate grading, provide real-time feedback, and reduce administrative burdens on teachers. This is highly relevant in EFL writing because teachers often spend a large amount of time correcting repeated grammar and mechanics errors. If AI tools can offer preliminary feedback on these matters, teachers can allocate their time to more meaningful pedagogical tasks, such as assisting students in developing arguments, improving organization, and reflecting on their revision decisions. In this way, AI does not remove the teacher's role, but redistributes the assessment workload.

Huang and Winkel (n.d.) explain that academic writing assessment is time-consuming and that AI-supported automated feedback offers benefits in terms of time efficiency, feedback consistency, and availability. Their discussion shows that AI tools can take over repetitive lower-order feedback tasks, thereby freeing teachers to focus on more complex pedagogical work. Similarly, Alnemrat et al. (2025) argue that AI-generated feedback may serve as a scalable complement to teacher feedback, especially in large, mixed-proficiency EFL classrooms.

AI-assisted marking also provides evidence for this point. Li et al. (2024) discovered that ChatGPT could evaluate written assignments, differentiate between various levels of assignment quality, and produce detailed rubric-based justifications. This indicates that AI can serve as a marking

assistant, especially when educators provide clear prompts and assessment criteria. However, the same study also shows that the quality of AI marking depends heavily on prompt design and task clarity. Therefore, AI may support assessment scalability, but it still requires teacher supervision to ensure that the feedback is fair, accurate, and pedagogically useful.

This is a significant opportunity because the value of formative assessment often depends on how consistently feedback can be delivered. If only a small number of students receive detailed, timely input, the pedagogical benefits remain uneven. AI can help reduce this inequity by making feedback more widely available. In practice, this means that teachers may be able to sustain formative assessment more realistically when AI handles routine feedback and teachers concentrate on more interpretive and strategic aspects of guidance.

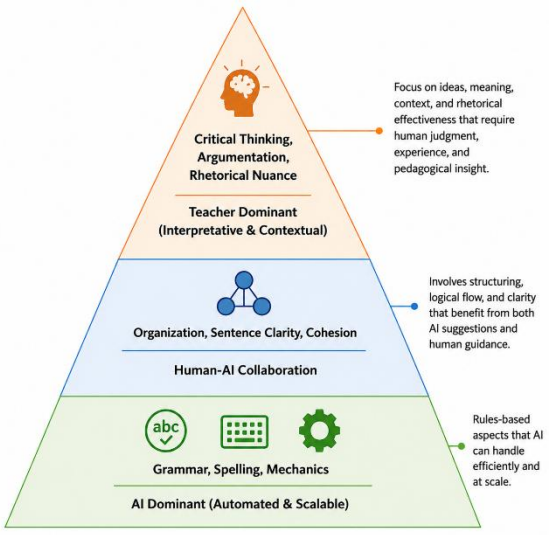
4. The major pedagogical challenge: AI is stronger on lower-order concerns than higher-order writing

Despite these advantages, the literature consistently shows that AI-powered tools are not equally effective across all aspects of writing. Their greatest strength lies in lower-order concerns such as grammar, mechanics, references, and surface-level coherence, while their ability to address higher-order concerns remains limited.

This limitation is evident in several studies. Alzahrani and Alotaibi (2024) found that ChatGPT improved several aspects of EFL writing, including task achievement, coherence and cohesion, lexical resource, and overall writing performance, but grammatical range and accuracy did not improve in the same way. This suggests that AI support may affect different writing components unevenly. Marzuki et al. (2023) also report that EFL teachers perceived AI writing tools as useful for improving content and organization, yet teacher interpretation remains necessary because content quality and organization depend on purpose, context, and audience. These findings suggest that AI can support writing improvement, but it cannot fully replace human assessment of complex writing quality.

Huang and Winkel (n.d.) clearly distinguish between lower-order concerns and higher-order concerns, noting that automated feedback is most useful for grammar, spelling, style, and reference-related issues, while teachers should focus on reasoning, argumentation, and content. Nazli et al. (2025) also report that although AI feedback improves grammar and coherence, it is less capable of handling contextual nuance and complex writing elements.

Figure 1. Hierarchy of Writing Assessment Concerns in AI-Assisted Pedagogy



This limitation is especially important in EFL writing because writing quality is not reducible to correctness alone. Higher-order writing features such as argument strength, logical development, audience awareness, rhetorical organization, and voice require interpretation that is often contextual, disciplinary, and pedagogically nuanced. Tahir et al. (2025) found that AI writing assistants could support writing, but their ability to foster critical thinking and deep reasoning remained uncertain. Lower-proficiency students in particular tended to

use AI for superficial adjustments rather than for more substantive engagement with ideas.

This means that AI-powered formative assessment should not be framed as a total assessment solution. Rather, it should be understood as partial support within a broader assessment ecology. Teachers remain essential for diagnosing higher-order problems, guiding rhetorical development, and interpreting student writing in context.

5. AI feedback must be connected to communicative and contextual meaning

Another important limitation is that AI tools may not fully understand communicative and contextual meaning. Williams (2025) argues that AI-based language tools may support grammar, vocabulary, pronunciation, and surface-level performance, but they are limited in developing communicative competence because they cannot fully replicate human interaction and sociolinguistic context. This point is relevant to EFL writing because writing is also a communicative activity. A grammatically correct text may still be weak if it does not address the audience, purpose, genre, or cultural context appropriately. Therefore, AI feedback should not be used only to make writing more accurate; it should be interpreted through teacher guidance so that students understand how their writing communicates meaning.

A similar concern appears in AI-based proficiency evaluation. Manggiasih et al. (2023) found that the SmallTalk2Me app offers objective and immediate evaluation, but it also has limitations in capturing the complexity of real-life communication and differences in speech patterns or accents. Although this study focuses on speaking proficiency, the implication also applies to writing assessment. AI systems may evaluate language based on predefined criteria, but real communication often involves nuance, intention, and context. In EFL writing, teachers are needed to help students understand whether their writing is effective for a particular reader and situation.

6. Overreliance on AI may weaken critical engagement with writing

A second pedagogical challenge concerns student overreliance on AI. Formative assessment is supposed to encourage active learning, self-reflection, and improvement. Yet if students begin to treat AI-generated suggestions as automatically correct, they may revise mechanically without understanding why changes are needed.

This concern is particularly important when AI tools are used for drafting and rewriting. Harunasari (2023) emphasizes that the responsible use of ChatGPT in EFL writing requires clear classroom policies and a record of students' questions or prompts. This is necessary because AI can easily move from being a feedback tool to becoming a text-generation tool. If students use AI to produce large parts of their writing without understanding the content, formative assessment loses its purpose. Instead of showing student development, the final text may show the ability of the AI system. Therefore, teachers need to make AI use transparent and connect it to reflection.

Tahir et al. (2025) explicitly warn that AI-powered writing assistants may encourage dependency, especially among less proficient learners who focus on surface-level editing rather than argumentation or reasoning. Nazli et al. (2025) similarly identify overreliance on AI as a concern in writing improvement. This is pedagogically significant because formative assessment should strengthen students' awareness of learning processes, not reduce them to compliance with automated suggestions.

In other words, AI may support revision but not necessarily reflection. If students are not taught how to evaluate AI feedback critically, then AI can weaken rather than

strengthen formative assessment. This is why pedagogical mediation is crucial. Students need guidance in deciding which AI suggestions to accept, reject, or question. Without such mediation, the formative function of AI feedback may be reduced to correction rather than learning.

7. Ethical issues in AI-powered formative assessment

Which, in fact is not without its own set of serious ethical concerns. According to Sheikh (2024), multiple factors such as data privacy, algorithmic bias and depersonalisation are the most serious obstacles in AI-based assessment. Ethical considerations have also been emphasized by Mousa (2025) referring specifically to issues about privacy, intellectual property and academic integrity and the adverse effect of AI on students critical thinking. Such issues are immediately pertinent to EFL writing because students could submit content that contains personal writing, drafts, ideas or academic assignments into AI tools. Now, this informal process of teachers encouraging AI use without any guidelines may lead to students not knowing whether they do not have a right to know how their data are utilized, or whether the corrections made by AI need to be cited.

Another big topic is academic integrity. In formative assessment, the process of learning is equal to or even more valuable than the final product. This is a problem if students submit writing that has merely been substantially generated or rewritten by AI, as it can make it difficult for teachers to interpret the student when assessing ability. It is due to this very reason that AI use needs to be recorded. Students can be asked to provide a statement explaining how they used AI, what specific tool they chose, the exact prompts they provided, the feedback they accepted and why and how they revised their writing. This does not do that by banning AI, but rather making its use visible and measurable. This also teaches students that responsible use of AI is part of the practice of writing in academia.

8. Teacher judgment remains central in AI-powered formative assessment

The evidence strongly indicates that the most impactful model is not AI vs teachers, but AI plus teachers. Alnemrat et al. Findings from the research of (2025) do not provide statistically significant difference between AI-generated comparison and teacher-generated feedback in regard to overall improvement in writing; however, their discussion leans strongly towards pedagogical supervision and ethics of deployment. Similarly, Valerio (2025) also reported superior effects when integrating AI-based tools with face-to-face instruction compared to either type of instruction alone.

Teacher judgment is essential too as the effectiveness of AI assessment relies on prompt quality, rubric, and task design. Li et al. This means that marking outcomes tend to vary depending on the prompts used for ChatGPT (2024). This means that AI is not neutral or unbiased, the feedback depends on instructions from human users. Hence, in EFL writing classrooms, teachers should plan clear prompts, rubrics and feedback procedures prior to asking students to engage with AI tools. Simply put. Students may receive feedback that is vague, incorrect or misaligned with the skills and standards taught in the classroom without teacher moderated guidance.

These findings underpin a hybrid formative assessment model. In this model, AI can offer real-time feedback on procedural and lower-order matters while putting teachers back in charge of decoding the complexity associated with higher-order writing issues (e.g., rubric construction or grading), providing nuanced responses to contextual and rhetorical demands, enabling reflection and revision techniques, permitting assessment-scoped WWCs to be pedagogically relevant.

Table 2. Division of Roles between AI and Human Educators in Writing Assessment

Assessment Aspect	*AI-Powered Tools Focus*	*Human Educator Focus*
Level of Concerns	**Lower-order concerns (LOCs)**: focus on technical accuracy.	**Higher-order concerns (HOCs)**: focus on meaning and strategy.
Technical Elements	Grammar, spelling, mechanics, and citation formats.	Argumentation, critical reasoning, and deep coherence.
Feedback Context	Automated, instant, and based on predefined linguistic patterns.	Strategic, interpretative, and based on communicative intent.
Scale and Timing	Providing frequent feedback during the drafting process.	Final judgment, mentoring, and ensuring pedagogical soundness.
Primary Goal	Reducing workload and improving assessment scalability.	Developing critical thinking and rhetorical voice.

This hybrid model is even more in line with international theories of formative assessment than either a wholly human or purely automated model. This is not to say that formative assessment is just more feedback, it is the right kinda feedback XTimingXto those who need in order to advance learning. AI is a solution in terms of timing and scale, but educators ultimately are still needed to provide depth, relevance, and pedagogical soundness.

9. Toward a balanced model of AI-powered formative assessment in EFL writing

The evidence strongly indicates that the most impactful model is not AI vs teachers, but AI plus teachers. Alnemrat et al. Findings from the research of (2025) do not provide statistically significant difference between AI-generated comparison and teacher-generated feedback in regard to overall improvement in writing; however, their discussion leans strongly towards pedagogical supervision and ethics of deployment. Similarly, Valerio (2025) also reported superior effects when integrating AI-based tools with face-to-face instruction compared to either type of instruction alone.

Teacher judgment is essential too as the effectiveness of AI assessment relies on prompt quality, rubric, and task design. Li et al. This means that marking outcomes tend to vary depending on the prompts used for ChatGPT (2024). This means that AI is not neutral or unbiased, the feedback depends on instructions from human users. Hence, in EFL writing classrooms, teachers should plan clear prompts, rubrics and feedback procedures prior to asking students to engage with AI tools. Simply put. Students may receive feedback that is vague, incorrect or misaligned with the skills and standards taught in the classroom without teacher moderated guidance.

These findings underpin a hybrid formative assessment model. In this model, AI can offer real-time feedback on procedural and lower-order matters while putting teachers back in charge of decoding the complexity associated with higher-order writing issues (e.g., rubric construction or grading), providing nuanced responses to contextual and rhetorical demands, enabling reflection and revision techniques, permitting assessment-scoped WWCs to be pedagogically relevant.

This hybrid model is even more in line with international theories of formative assessment than either a wholly human or purely automated model. This is not to say that formative assessment is just more

feedback, it is the right kinda feedback XTimingXto those who need in order to advance learning. AI is a solution in terms of timing and scale, but educators ultimately are still needed to provide depth, relevance, and pedagogical soundness.

Conclusion

EFL writing with AI-powered tools promises quality pedagogical opportunities for formative assessment. The literature review of this paper shows that these tools can provide timely feedback and revision support, facilitate self-regulated learning, and strengthen the scalability of writing feedback in settings where teacher time is limited. Studies by Nazli et al. (2025), Huang and Winkel (n.d.), a Williyan e a. 2017; Yates 2025) describe, AI can enhance writing pedagogy when it is embedded into iterative feedback and revision practices.

Simultaneously on the other hand, this paper points to key pedagogical issues. Beyond the mechanics of writing like grammar and syntax, AI has been ahead in almost all lower-order concerns than higher-order ones like reasoning, rhetorical quality or argument development. Moreover, real dependency on AI might undermine students' significant engagement with revision in that they may accept those suggestions at face value. Tahir et al. (2025) and Alnemrat et al. (2025) and argue for the continual need for human pedagogical supervision, especially when quality of writing hinges on contextualized, esthetic interpretation.

The reviews of studies in this paper provide strong but also limited evidence that AI is a powerful tool. Research on the improvements in writing through AI writing tools (e.g., grammar, vocabulary, organization, motivation and performance) has been reported (Al Mahmud, 2023; Alzahrani & Alotaibi, 2024; Song & Song, 2023; Wale & Kassahun, 2024). Research into how AI is assessed indicates that automated feedback and AI-assisted marking aspects of less workload alongside scalable feedback, contingent upon careful prompt. ethical guidelines and supervising teachers (Li et al., 2024; Mousa, 2025; Sheikh, 2024). Lastly, research on AI-supported language assessment and communicative competence also cautions that AI cannot unpack context, genuinely human interaction and meaningful complexity (Manggiasih et al., 2023; Williams, 2025). Hence there is a necessity for cautious yet positive evolution of AI based formative assessment.

For these reasons, AI-powered formative assessment should be understood as a hybrid pedagogical model in which AI supports routine and immediate feedback, while teachers remain central in guiding higher-order writing development and ensuring meaningful learning. In this sense, the most productive future for AI in EFL writing assessment is not full automation, but pedagogically balanced integration.

Figure 2. The Proposed Hybrid Formative Assessment Cycle

In practical terms, AI-powered formative assessment should be implemented gradually and reflectively. Teachers can begin by using AI for low-stakes drafting activities, grammar feedback, vocabulary suggestions, and revision logs before applying it to more complex writing assessment tasks. This gradual integration

allows students to develop confidence and AI literacy while helping teachers observe how AI feedback influences students' writing behavior and revision quality.

Future EFL writing pedagogy should thus concentrate not only on the integration of AI tools but also on fostering AI literacy among both educators and learners. Educators require training in crafting AI-assisted writing assignments, assessing AI-generated feedback, and formulating ethical policies for the classroom. Learners need support in utilizing AI with a critical perspective, recording their revision processes, and recognizing that AI feedback serves as a learning resource rather than a replacement for independent thought. When these prerequisites are fulfilled, AI-enhanced formative assessment can evolve into a significant component of EFL writing education.

References

Alnemrat, A., Aldamen, H., Almashour, M., Al-Deaibes, M., & AlSharefeen, R. (2025). AI vs. teacher feedback on EFL argumentative writing: A quantitative study. *Frontiers in Education, 10*, Article 1614673.

Alzahrani, F. K. J., & Alotaibi, H. H. (2024). The impact of artificial intelligence on enhancing EFL writing skills among high school students. *Journal of Educational and Human Sciences*, *(34)*, 226–240. https://doi.org/10.33193/JEAHS.34.2024.482

Aysu, S. (2024). The use of AI-powered tools in language learning. In S. Aysu & H. Göçerler (Eds.), *Dil Eğitimi ve Öğretiminde Dijital Yönelimler* (pp. 1–18). Özgür Publications. https://doi.org/10.58830/ozgur.pub476

Harunasari, S. Y. (2023). Examining the effectiveness of AI-integrated approach in EFL writing: A case of ChatGPT. *International Journal of Progressive Sciences and Technologies (IJPSAT)*, *39*(2), 357–368. https://doi.org/10.52155/ijpsat.v39.2.5516

Hasibuan, R., Parta, I. B. M. W., Sholihah, H. I., Damayanto, A., & Farihatun. (2023). Transformation of Indonesian language learning with artificial intelligence applications: The era of the independent curriculum for learning in universities in Indonesia. *Indonesian Journal of Education (INJOE)*, *3*(2), 341–363.

Huang, Z., & Winkel, W. (n.d.). *Automated feedback: An AI-powered tool to scale micro-level feedback for better academic writing*.

Hussain, S. M., Satti, S. M. J., Khan, Z., & Um-e-Safia. (2024). AI-powered personalized learning: Advancing language education in the digital era. *Journal of Social Signs Review*, *2*(4), 730–740.

Li, J., Jangamreddy, N., Bhansali, R., Hisamoto, R., Zaphir, L., Dyda, A., & Glencross, M. (2024). AI-assisted marking: Functionality and limitations of ChatGPT in written assessment evaluation. *Australasian Journal of Educational Technology*, *40*(4), 56–72. https://doi.org/10.14742/ajet.9463

Mahmud, F. A. (2023). Investigating EFL students' writing skills through artificial intelligence: Wordtune application as a tool. *Journal of Language Teaching and Research*, *14*(5), 1395–1404. https://doi.org/10.17507/jltr.1405.28

Manggiasih, L. A., Loreana, Y. R., Azizah, A., & Nurjati, N. (2023). Strengths and limitations of SmallTalk2Me app in English language proficiency evaluation. *TELL: Teaching of English Language and Literature Journal*, *11*(2), 134–150.

https://doi.org/10.30651/tell.v11i2.19560

Marzuki, Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*, *10*(2), Article 2236469. https://doi.org/10.1080/2331186X.2023.2236469

Mousa, M. (2025). AI-supported formative assessments: Enhancing student-centered learning and teacher perceptions. *Journal of Pedagogy and Education Science (JPES)*, *4*(2), 127–141. https://doi.org/10.56741/jpes.v4i02.857

Nazli, K., Jumani, S., & Masum, R. (2025). The effectiveness of AI-powered corrective feedback on students' writing performance. *Pakistan Social Sciences Review,* *9*(1), 391–403. https://doi.org/10.35484/pssr.2025(9-I)31

Nurhayati, A. (2020). The implementation of formative assessment in EFL writing: A case study at a secondary school in Indonesia. *Pedagogy: Journal of English Language Teaching, 8*(2), 126–137. https://doi.org/10.32332/pedagogy.v8i2.2263

Parsa, M., Alam, M. R., & Mihailidis, A. (2021). Towards AI-powered language assessment tools. *Research Square*. https://doi.org/10.21203/rs.3.rs-246079/v1

Sheikh, F. (2024). AI-powered assessment and feedback: Enhancing student engagement and performance. *AI Edify Journal*, *1*(4), 30–37.

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, *14*, Article 1260843. https://doi.org/10.3389/fpsyg.2023.1260843

Tahir, M., Jahrir, A. S., Asrifan, A., & Ariatna, A. (2025). Examining the role of AI-powered writing assistants in enhancing critical thinking in EFL academic writing. *JOLLT Journal of Languages and Language Teaching, 13*(4), 1817–1831. https://doi.org/10.33394/jollt.v13i4.15077

Valerio, M. T. B. (2025). Enhancing argumentative essay skills: Evaluating the efficacy of AI-powered tools versus the traditional approach among learners. *Arab World English Journal (AWEJ) Special Issue on Artificial Intelligence*, 270–289. https://doi.org/10.24093/awej/AI.15

Wale, B. D., & Kassahun, Y. F. (2024). The transformative power of AI writing technologies: Enhancing EFL writing instruction through the integrative use of Writerly and Google Docs. *Human Behavior and Emerging Technologies*, *2024*, Article 9221377. https://doi.org/10.1155/2024/9221377

Wijaya, K. F. (2022). Indonesian EFL teachers' perceptions on formative assessment in writing. *Journal of English Teaching,* *8*(1), 59–70. https://doi.org/10.33541/jet.v8i1.3359

Williams, B. (2025). Assessing the limitations of AI-based language tools in enhancing communicative competence in higher secondary classrooms. *ResearchGate*.

Williyan, A., Ilyas, M., Shahat, S., Guntur, M., & Rosalina, U. (2025). AI-powered pedagogy for argumentative essays in tertiary ELT: Insights from focus groups. *Journal of General Education and Humanities,* *4*(3), 671–686. https://doi.org/10.58421/gehu.v4i3.423