

Generative AI in Student Essays: English Teachers' Perspectives on Effective Assessment Methods

Valentyna Lukianenko – Zoia Kornieva

DOI: 10.18355/XL.2024.17.04.14

Abstract

The increasing use of AI-assisted writing tools in education presents new challenges in ensuring the integrity and originality of student essays. This study investigates English teachers' perceptions of various essay assessment methods in light of these challenges. Through a mixed-methods approach, combining quantitative Likert-scale data and qualitative open-ended responses, the research reflects the perspectives of 50 experienced educators from the Faculty of Linguistics at the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute." The study investigates teachers' familiarity with AI tools, the challenges they face in assessing AI-assisted student work, and their views on AI's role in aiding or potentially hindering student writing skills. Findings indicate that while many educators support AI for tasks like grammar and proofreading, concerns persist over its impact on students' critical thinking and originality. Assessment methods currently in use include in-class handwritten essays, oral defenses, AI-detection software, and portfolios. The results indicate that hybrid assessment methods are rated as the most effective for evaluating AI-influenced essays, indicating a preference for combining traditional and AI-specific techniques. Implications and recommendations for research and practice have been outlined, emphasizing a balanced approach that promotes ethical AI usage, critical thinking, and academic integrity in writing assessments.

Key words: AI in education, essay assessment methods, AI-assisted writing, academic integrity, hybrid assessment approaches, AI detection tools

Introduction

Incorporating artificial intelligence (AI) tools in education has introduced advantages and challenges, particularly in evaluating student-written works. Although AI can enhance productivity and creativity, it also complicates the assessment of originality, authenticity, and depth in student assignments. Essay assignments are essential in English language acquisition, enhancing linguistic proficiency, critical thinking, and effective communication. Writing enables students to exhibit their proficiency in language and vocabulary and their capacity to develop arguments, incorporate research, and show creativity. The increasing prevalence of generative AI presents significant challenges to conventional assessment methods. This study investigates English teachers' perceptions of various essay evaluation methods, considering AI's impact on writing, to evaluate the efficacy of these methods in maintaining the integrity of the essay-writing process.

Literature Review

Teacher and Student Perceptions of Generative AI in Writing and Learning

Generative AI tools like ChatGPT are increasingly integrated into education, leading to mixed reactions from students and teachers about their role in learning and assessment. For students, these tools offer support in areas like grammar and structure while raising concerns about creativity and originality. Tossell et al. (2024) found that students initially viewed AI as a "cheating tool" but gradually saw its potential as a collaborative aid for idea generation, with a preference for teacher oversight in grading and assessment. On the other hand, teachers express mixed feelings,

appreciating AI's potential for scaffolding but concerned about its impact on their instructional role and transparency (Soochwan & Song, 2024; Kim & Kim, 2022). Malik et al. (2023) emphasize the importance of using AI as a support rather than a replacement for human teaching, ensuring that student learning remains central. Kizilcec et al. (2024) broaden this idea by comparing educators' and students' attitudes in three countries, revealing that educators are more inclined to adopt AI for assessments emphasizing critical thinking, while students worry AI may diminish creativity. Together, these studies suggest that while AI offers valuable support for both students and teachers, its integration requires careful balance and strategic adaptation to preserve academic integrity and foster higher-order thinking.

Impact of AI on Assessment and Student Performance

The integration of AI in academic assessment has caused educators to rethink traditional grading methods, particularly concerning the authenticity of student work. AI's influence on student performance is a growing area of research, with a focus on how assessment design can address issues such as academic dishonesty. Smerdon (2024) found that students primarily used AI as a supplemental tool in research proposals, which did not negatively impact their academic performance. This finding suggests that, when used responsibly, AI can enhance learning outcomes, offering support in research without undermining academic integrity.

Similarly, Guo et al. (2024) examined AI-assisted peer feedback and found that it significantly improved writing quality among EFL students. Their study underscores AI's potential to provide meaningful support in feedback-intensive environments where human resources may be limited, suggesting a practical role for AI in enhancing educational processes. However, concerns about AI's role in academic dishonesty persist. Sweeney (2023) explored the issue of AI-driven cheating within a UK business school, particularly in the context of essay mills. The study argued for a more intentional assessment design and providing student-centered feedback to counter academic dishonesty. The researcher highlighted the need for alternative assessment methods to better prepare students for the workforce.

These studies collectively suggest that while AI can enhance educational outcomes, its integration must be carefully managed to avoid promoting academic dishonesty. By incorporating ethical support and exploring alternative forms of assessment, educators can ensure that AI is used to promote genuine student learning and support academic integrity.

Challenges in Detecting AI-Generated Content

The rise of generative AI has significantly complicated the task of distinguishing between human and AI-generated work, raising concerns about academic integrity. Research in this area focuses on the detectability of AI-generated content and the limitations of current detection technologies. Fleckenstein et al. (2024) found that even experienced educators struggled to accurately identify AI-generated content, demonstrating that traditional assessment methods are insufficient for detecting AI use. Waltzer et al. (2023) further supported these findings, showing that both teachers and students had difficulty distinguishing AI-written work from human-generated content. This emphasizes the urgent need for more effective detection tools and assessment methods to safeguard the authenticity of academic assessments in an increasingly AI-driven environment.

As generative AI advances, traditional plagiarism detection tools struggle to keep up with its evolving capabilities. Santra and Majhi (2023) addressed the limitations of current plagiarism detection tools, which often cannot recognize AI-generated content due to its distinct syntactical features. Their study emphasized the need for AI-based detection tools to adapt to these unique characteristics. By focusing on developing detection methods that account for the complexities of machine-generated text, their research highlights a critical area for improvement in maintaining academic integrity.

Krishna et al. (2023) took a different approach by exploring retrieval-based defenses against AI paraphrasing, offering a promising solution to the detection problem. These tools compare suspicious text with a database of known AI-generated content. Their system can identify patterns to identify AI authorship by retrieving semantically similar AI-generated texts, even after paraphrasing. This method suggests that focusing on retrieval-based tools can help institutions keep pace with AI's growing sophistication and more effectively detect AI-generated content.

These studies collectively point to the emerging challenge of reliably detecting AI-generated work. To uphold academic integrity, educators must adopt advanced detection methods and explore alternative strategies for verifying student-authored work in the context of AI's rapid development.

Ethical and Practical Considerations of AI's impact in ESL contexts

The ethical implications of integrating AI in higher education center around fairness, transparency, and academic integrity. To address these concerns, recent research has proposed various strategies to ensure responsible AI use, all of which share a common goal: preserving academic standards while adapting to new technological realities. Sharples (2022) raised the concern that AI could facilitate cheating by allowing students to complete assignments with minimal effort. Sharples recommended developing AI tools that encourage genuine learning to counter this risk. These tools would maintain academic rigor while still supporting students' use of AI in ways that foster more profound engagement with the material. Developing this idea, Fyfe (2023) proposed a practical solution by encouraging students to combine AI-generated text with their own writing. This approach improved students' understanding of academic honesty and underlined the importance of establishing clear guidelines on the acceptable use of AI. This study highlights that practical exercises combining human input with AI assistance can teach responsible AI use.

Sweeney (2023) continued the research on ethical AI integration and stressed the importance of transparency in assessments, suggesting that clear communication around AI's role in evaluation can help combat the rise of essay mills and protect academic integrity. In a related work, Alexander et al. (2023) explored AI's impact in ESL contexts. They noted that AI-generated work often exhibited unrealistic levels of linguistic accuracy, raising ethical concerns about the fairness of its use in language learning. To address this, the authors advocated for digital literacy programs and policy reforms, helping educators navigate AI's ethical challenges while ensuring that students are assessed fairly.

The reviewed literature reveals a complex relationship between generative AI and higher education, balancing the benefits of AI-enhanced learning with challenges in maintaining academic integrity and ethical standards. Students and educators see the value of AI as a supportive tool, yet concerns about its potential to undermine creativity, fairness, and authenticity persist. Despite the growing body of literature on AI in education, there is limited empirical research exploring how educators perceive the effectiveness of various assessment methods in the context of AI-assisted writing. Most studies focus on the technological capabilities of AI detection tools, but few address the pedagogical strategies that educators find most effective in maintaining the integrity of student essays. Additionally, research has yet to examine how educators view hybrid methods that combine traditional approaches with AI tools for more robust assessment. This study aims to address these gaps by providing quantitative and qualitative data on educators' perceptions of various essay assessment methods, specifically concerning AI-assisted student writing.

Research Questions

1. How familiar are English teachers with AI tools used by students for writing?

2. What are English teachers' primary challenges in assessing the authenticity and originality of student essays influenced by AI tools?
3. How do teachers perceive using AI tools for partial assistance (e.g., grammar checks)?
4. Which assessment methods do teachers currently use or consider using to evaluate student essays in light of potential AI usage, and how effective do they find these methods?

Methods

This study employed a mixed-methods, survey-based design to examine teachers' perceptions of various assessment methods for student essays in the context of potential AI tool usage. It incorporated both quantitative Likert-scale items and qualitative open-ended responses.

Participants

The research participants (N=50) were educational staff of the Faculty of Linguistics, the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute." All educators, whether teaching linguistic students or engineering students, were involved in assessing student written works (essays, papers, conference abstracts, course papers, etc), selected for their experience in traditional and AI-related assessment methods.

Data Collection

The data were collected through an online questionnaire distributed via Google Forms. The questionnaire was designed to investigate English teachers' perspectives on writing assessment in the age of AI. It comprised 15 questions divided into three main sections, each addressing different aspects of the topic. Section 1 included four questions on demographics: age, teaching experience, and student demographics. Section 2 contained five questions assessing teachers' familiarity with AI tools, experiences with AI-influenced student work, and challenges in evaluating such work. Section 3, with six questions, focused on current and prospective assessment methods for AI-influenced essays. The questionnaire included Likert-scale, multiple-choice, and open-ended questions, allowing for a comprehensive analysis of teacher perspectives.

The questionnaire's validity is strengthened by its strategic design, which comprehensively addresses key issues related to AI in student essay writing. Including closed-ended and open-ended questions enables a balanced collection of quantitative and qualitative data, enhancing the depth and reliability of insights gathered.

Data Analysis Tools

Percentages were calculated for each Likert scale response to understand the distribution of perceptions regarding each assessment method. We calculated the mean, median, and standard deviation for each Likert scale response to provide a detailed understanding of teachers' perceptions regarding various assessment methods. SPSS (Statistical Package for the Social Sciences) was used to manage data and support further statistical analysis.

We conducted an analysis to explore the relationship between respondents' familiarity with AI tools and their perceived challenges in assessing AI-influenced student work. Given the categorical nature of the data, we calculated weighted averages for each variable and examined potential relationships between them.

To analyze the open-ended responses on teachers' attitudes towards students' use of AI for partial assistance, we conducted a thematic analysis by following several systematic steps: first, responses were coded; next, major themes were identified through constant comparison, grouping similar responses while highlighting unique insights; finally, these themes were interpreted to reveal common attitudes and

concerns about AI use in language education, illustrated with direct quotations to emphasize key points.

These tools allowed us to systematically analyze both the quantitative and qualitative data, providing a comprehensive understanding of the survey responses.

Ethical Issues

The study adhered to ethical research standards, ensuring informed consent was obtained from all participants prior to their involvement. Participants were informed about the voluntary nature of the survey and their right to withdraw at any point. The study design was reviewed for compliance with ethical standards in educational research, and no participants were exposed to any harm or discomfort during the process.

Results and Discussion

The first part of the questionnaire included demographic patterns. Respondents are fairly distributed across age groups, with the 30-39 and 40-49 age ranges each representing 30% of the sample. The 50-59 age group comprises 26% of the population, indicating a substantial portion of mid-career educators. Additionally, 12% of respondents are aged 60 and above, while 2% are in the 20-29 age range.

88% of respondents teach English to Engineering and IT students, 14% to Natural Sciences students, 12% to Linguistics students, and 8% to Social Sciences students, suggesting some diversity in student fields but a predominance of technically oriented education.

More than half of all respondents (52%) have over 21 years of experience teaching English, 22% have 11-15 years of experience, and 18% have been teaching for 16-20 years. A smaller percentage, 6%, have 6-10 years of experience, and only 2% have taught English for 0-5 years.

RQ1: 1. How familiar are English teachers with AI tools used by students for writing?

The second part of the questionnaire was devoted to current assessment challenges (see Fig.1). Regarding the question about familiarity with AI tools for essay writing, the majority of respondents (86%) indicated that they are familiar or somewhat familiar with AI tools used by students. Only 12% are neutral, and 2% are unfamiliar. This suggests that most educators have a basic understanding of AI tools in the student context.

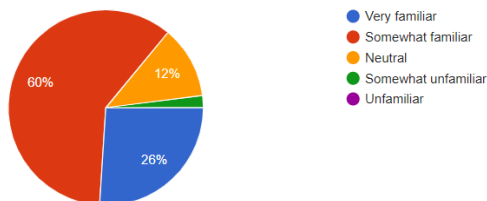


Figure 1: English teachers' familiarity with AI tools used by students for essay writing

Answering the question, "Have you encountered essays or other students' written works that you suspect were influenced or generated by AI tools?" an overwhelming 80% of respondents have encountered, or suspect they have encountered, AI-generated student work. 14% of respondents reported finding it hard to say whether AI had been involved, and 7% have not encountered AI-generated essays.

RQ2: What are English teachers' primary challenges in assessing the authenticity and originality of student essays influenced by AI tools?

Assessing the authenticity of AI-assisted writing is a significant concern, with 60% of respondents finding it somewhat challenging and 20% considering it very challenging. One teacher noted, *“It is acceptable to use AI to generate ideas or insights. However, if students submit it as their genuine work for grades, it breaches academic integrity.”* Few respondents find it less challenging, but our survey results show that verifying AI involvement is a prominent issue (see Fig.2).

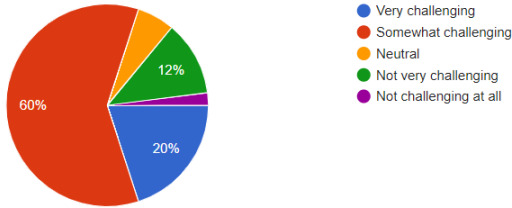


Figure 2: Perceived challenges in assessing authenticity and originality of AI-influenced essays

Figure 3 illustrates several key challenges teachers encounter when assessing essays potentially influenced by AI. One of the most prevalent concerns (54%) is the difficulty in determining the authenticity of student work and maintaining fairness in grading (50%) when AI use is suspected. One respondent mentioned, *“For the students nowadays, it is inevitable to use AI tools. Indeed, it contradicts my vision of a modern, independent, critically-thinking future specialist.”* Other educators expressed the need for clearer guidelines: *“I believe that AI tools can simplify the process of acquiring a foreign language if used correctly. Unfortunately, there are quite a lot of students who use them for cheating.”* In addition, nearly 40% of respondents express uncertainty about the extent and type of AI assistance students might have employed, further complicating the assessment process. A lack of clear guidelines also complicates teachers' ability to assess AI-assisted work effectively. Detecting AI-generated content poses a challenge, as many teachers find it difficult to identify when and how AI tools have been employed. Interestingly, a few respondents indicated they have not yet encountered significant difficulties in this area. However, a few mentioned resource limitations, particularly concerning comprehensive AI detection software availability.

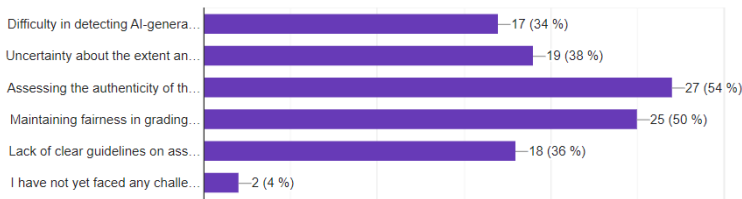


Figure 3. Challenges faced by educators in assessing essays that might potentially be influenced by AI

The analysis of teachers' familiarity with AI tools ($M = 3.10$) and their perceived challenges in assessing AI-influenced student work ($M = 3.84$) reveals some interesting results. Teachers reported a relatively high level of familiarity with AI,

with the majority indicating that they are either "very familiar" or "somewhat familiar" with such tools. At the same time, most of them found assessing AI-influenced work somewhat challenging, as reflected in the higher average score for perceived challenges. These weighted averages suggest a potential relationship: as familiarity with AI tools increases, so do the perceived challenges in assessing student work influenced by AI. This may indicate that the more familiar teachers become with AI, the more aware they are of the complexities of maintaining academic integrity when assessing such work.

RQ3. How do teachers perceive using AI tools for partial assistance (e.g., grammar checks)?

To answer the RQ3, we performed a thematic analysis of responses to the open question on students using AI tools for partial assistance, like proofreading and grammar checks. This revealed diverse perspectives among educators. We have outlined the following themes.

1. Support for partial AI assistance. A substantial number of English teachers view AI tools positively for grammar and proofreading support, emphasizing that these tools can benefit students' writing skills by offering immediate feedback. One respondent noted, *"I support this practice. It helps to enhance learners' writing abilities and improve overall writing proficiency by analyzing grammar and structure of the text."* Another teacher said, *"Using AI tools for partial assistance, such as proofreading and grammar checks, can be a beneficial practice: instant feedback to students, helping them identify and correct errors, improving writing skills."* These responses reflect a belief that AI tools can play a valuable role in language development when used responsibly.

2. Conditional support depending on academic level or task type. Some teachers believe that the appropriateness of AI assistance depends on the student's level and the educational objectives. For example, one educator stated, *"I think that partial assistance like proofreading may be allowed at the academic level—when Master or PhD students write papers, abstracts, etc."* Another response suggested that *"At the undergraduate level, students shouldn't use any AI assistance in writing as teachers should see their mistakes and teach students to improve them,"* reflecting a view that AI use should be limited based on the learning stage and goal.

3. Concerns about overreliance on AI. A number of educators are concerned that AI tools might lead to overreliance, potentially weakening students' critical thinking and skill development. One respondent emphasized, *"Relying on these tools can lead to students becoming less engaged in the learning process, as they may stop thinking critically and developing their own skills. This practice undermines the goal of education, which is to encourage independent thought, problem-solving abilities, and a deep understanding of the material."* Another educator noted, *"In my opinion, using AI by students for partial assistance, such as proofreading or grammar checks, may be a normal practice in a digital age. Although students mostly overuse it and it is very difficult to control this process."* This highlights a concern that students may misuse AI tools, which could compromise their learning experience.

4. AI Literacy and ethical use. Many educators emphasized the need to educate students on responsible and ethical AI use. One respondent stated, *"I believe AI tools can be used in the learning process for self-assessment and mistake analysis. However, students should be taught to uphold academic integrity and avoid using such tools when taking tests."* Another teacher added, *"I support [AI] but we (teachers and students) must be taught [how to use it]."* Meanwhile, another respondent expressed, *"It is acceptable to use AI to generate ideas or insights. If the essay or other writing material was generated by AI, and students submit it as their genuine work to get grades, then it breaches the academic integrity standards."* These

reflections underscore the importance of developing AI literacy, ensuring that students understand when and how AI use is appropriate in academic settings.

The findings reveal that teachers generally support using AI tools in a supplementary role, particularly for tasks such as proofreading and grammar checks, which aligns with Putra’s (2023) observations that AI writing tools play a key role in detecting and correcting errors efficiently. As one respondent noted, “*To strike a balance is very important,*” emphasizing the need for a balanced approach where AI aids, rather than replaces, the learning process. While many believe that AI can enhance learning when used responsibly, there are concerns about overreliance and its impact on students’ independent learning skills.

RQ4: Which assessment methods do teachers currently use or consider using to evaluate student essays in light of potential AI usage, and how effective do they find these methods?

The data on current and future assessment methods used by educators (see Fig.4) reveal distinct patterns in adapting to AI’s impact on student essay writing. Educators are actively using a mix of methods, with a strong emphasis on strategies that promote authenticity and originality, and they also show significant interest in integrating additional approaches in the future. Table 1 presents the effectiveness ratings of various assessment methods for AI-influenced student essays based on teacher responses.

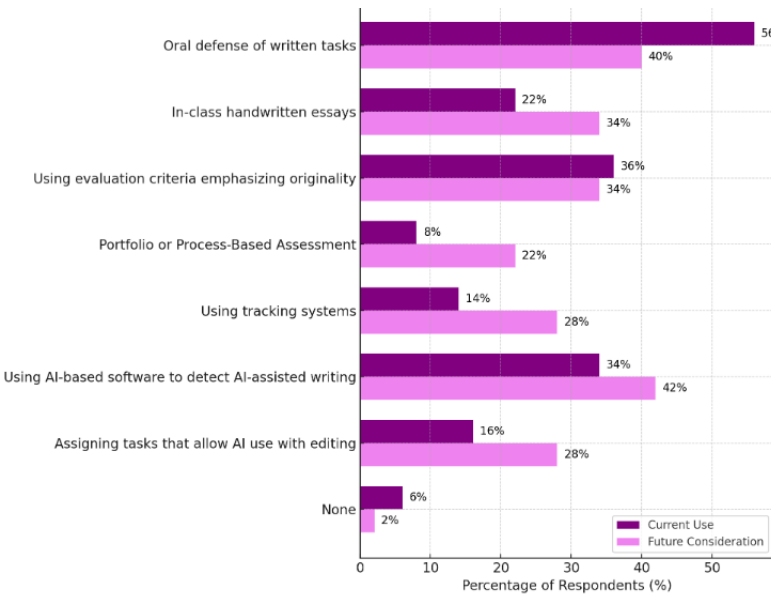


Figure 4. Current and future considerations on methods for assessing student essays influenced by AI

One of the most trusted traditional methods for ensuring authenticity in student writing is the **in-class handwritten essay**. Limiting access to external resources, including AI tools, offers immediate control over the writing environment and helps ensure that students’ work is entirely their own. In our study, 56% of respondents currently use this method, with 40% planning to implement it in the future. While it guarantees short-term authenticity, there are concerns about its practicality.

Handwritten tasks may not allow for the same depth of research or refinement that longer-term assignments can achieve. However, in-class handwritten essays remain popular among educators, with 78% finding them “Very Effective” and “Somewhat Effective.” Teachers value this method for preventing students from accessing AI tools during the task, although a small minority (8%) rated it as somewhat ineffective, possibly due to practical challenges or the limited feedback it allows. With a mean score of 3.92 and a standard deviation of 0.82, in-class essays are considered useful but not universally preferred.

Table 1. Effectiveness ratings of assessment methods for AI-influenced student essays

Method	5 (%)	4 (%)	3 (%)	2 (%)	1 (%)	Mean	Median	SD
Oral defense of written tasks	28	52	14	4	2	4.00	4.0	0.87
In-class handwritten essays	22	56	14	8	-	3.92	4.0	0.82
Evaluation criteria for originality	40	40	16	4	-	4.16	4.0	0.83
Portfolio or Process-Based Assessment	14	46	34	4	2	3.66	4.0	0.84
Tracking systems (e.g., Google Docs)	14	56	24	4	2	3.76	4.0	0.81
AI-based software to detect AI writing	20	40	32	8	-	3.72	4.0	0.87
Tasks allowing AI-assisted editing	28	48	24	-	-	4.04	4.0	0.72
Hybrid methods	52	38	10	-	-	4.42	5.0	0.67

*5 -Very Effective (%); 4- Somewhat Effective (%); 3-Neutral (%); 2-Somewhat Ineffective (%); 1-Very Ineffective (%)

Recent literature further supports using in-class handwritten essays to mitigate AI influence. Plate et al. (2023) investigated how traditional methods, like handwritten essays, offer educators a reliable means of evaluating genuine student effort in an AI-augmented environment. Sullivan et al. (2023) highlighted the need to update examination formats, such as shifting toward oral exams and in-person assessments, to reduce the risks of AI misuse, academic dishonesty, and over-reliance on AI and address concerns about academic integrity while logistical challenges exist. As noted by a minority of respondents in our study, handwritten tasks ensure that student assessments reflect their true capabilities and are free from outside assistance. This combination of tradition and control makes in-class handwritten essays a valuable, though not perfect, tool for addressing the growing presence of AI in education.

Oral defense of written tasks is another method used to verify students' understanding and authorship of their work. Teachers can evaluate whether students know their submitted content by asking them to explain their essays or discuss their writing choices. However, while this approach may help confirm the student's authorship, it is time-consuming and may not be feasible in large classes or institutions with limited resources. Additionally, it does not completely eliminate the possibility that AI was used to generate the written work.

The oral defense of written tasks is widely regarded as a highly effective method for assessing student essays, with 80% of respondents rating it as “Very Effective” and “Somewhat Effective.” Teachers appreciate this approach as it involves students actively participating in the writing process, encouraging deeper involvement in their own work. This method’s mean score of 4.00 underscores its value in ensuring student understanding and minimizing AI misuse in writing.

Currently, oral defenses of written tasks are the most frequently used method, with 56% of educators implementing them to verify the authenticity of student work. This approach emphasizes personal accountability, as one educator explained, “*I promise students to assess their written work only after they have retold the content of their compositions.*” For future use, 40% of respondents plan to continue or expand the use of oral defenses. This approach reinforces student responsibility and aligns with educators’ goals to verify understanding and uphold academic integrity in the writing process.

In addition to oral defenses and handwritten essays, 36% of educators currently use **evaluation criteria emphasizing originality, critical analysis, and personal insight**—qualities that are harder for AI to replicate. These criteria are valued as effective tools for assessing independent thinking, with 34% of educators considering this approach for future assessments, suggesting confidence in its effectiveness beyond what AI-generated content can easily achieve.

Originality-focused evaluation criteria are widely regarded as an effective approach for assessing student essays, with 40% of educators rating it as “Very Effective” and another 40% as “Somewhat Effective.” These criteria emphasize critical analysis, originality, and personal insight—elements AI-generated content often lacks. The mean score of 4.16 and a relatively low standard deviation of 0.83 highlight educators’ strong confidence in this approach as a reliable measure for promoting independent thinking and reducing reliance on AI assistance. This aligns closely with the concept of “voice” in written language, as discussed by Matsuda and Jeffery (2012). Voice captures a writer’s unique identity and perspective, which can be challenging for AI to replicate, especially in contexts requiring originality and critical thought.

As AI tools like ChatGPT become more advanced, the development of detection tools is evolving. **AI-based detection software**, such as Turnitin’s AI detection feature or ZeroGPT, is increasingly popular among educators who are concerned about academic integrity. These tools use algorithms to analyze text and estimate the likelihood that AI-generated it, offering a verification layer that traditional plagiarism detection software cannot provide. However, AI detection faces significant limitations. As AI continues to improve its ability to imitate human writing styles, detection software struggles to distinguish between AI-generated and human-authored text reliably. This is further complicated by AI tools to proofread their work. Although such editing tools aim to enhance grammar and clarity, the software may still detect their use, raising concerns about fairness.

Using AI detection software is gradually becoming part of standard practice, with 34% of English teachers in this study currently relying on it as an essential tool. One respondent described the challenges: “*I use programs that check for AI involvement, but it’s still hard to be certain.*” 42% of respondents plan to use AI-detection software in the future, indicating a trend toward integrating technology to help monitor student writing. As one educator emphasized, “*We are not only teaching language but also learning to use new technologies to make our work easier,*” highlighting the dual role of educators as both instructors and users of innovative tools.

Despite its growing adoption, AI detection software receives mixed evaluations from educators. Only 20% found these tools to be “Very Effective,” while 32% rated them neutrally, reflecting a mix of trust and skepticism. Concerns over reliability and accuracy are evident in the mean score of 3.72 and a relatively high standard deviation

of 0.87. Although considered somewhat helpful, these tools are not universally accepted as a definitive solution.

The findings from our study highlight the challenges that AI-generated content poses for educators, which is supported by literature (Fiialka et al., 2023; Weber-Wulff et al., 2023). After analyzing extensive research, Farrelly and Baker (2023) conclude that most current AI detection tools are generally unreliable. Their work suggests that, despite advancements, educators face uncertainty regarding the scope and nature of AI assistance students may use. Researchers observed that even established platforms often report high originality scores for AI-generated content, complicating efforts to identify and address academic misconduct. The study by Lukianenko et al. (2024) supports these findings by emphasizing the limitations of AI detection tools, particularly when content is paraphrased, and highlighting the importance of promoting ethical AI use in education. Despite their limitations, the growing reliance on these tools suggests the need for more comprehensive assessment methods that do not solely depend on automated detection systems.

Portfolio assessments, which involve collecting student work over a period of time, allow teachers to evaluate a student's growth and development of writing abilities through drafts, revisions, and final submissions. Portfolios can provide a comprehensive view of students' understanding of language, ideas, and their ability to reflect on their learning process. They encourage self-assessment, critical thinking, and the development of autonomy by allowing students to document their progress over time (Cimermanová & Jelavić, 2024). However, even portfolios may not guarantee authenticity in an age of AI. Students can easily incorporate AI-generated content into their work at any stage of the writing process, making it challenging for teachers to differentiate between student-created and AI-assisted writing.

Portfolio is a less common assessment method among our respondents, with only 8% of educators currently using it. However, 14% of respondents rated it as "Very Effective," and 34% viewed it neutrally. The response variability, reflected by a standard deviation of 0.84, indicates divided opinions on its effectiveness in detecting AI presence. The mean score of 3.66 suggests that while some educators appreciate this long-term approach, others may find it less effective for addressing AI-related challenges. Despite these challenges, portfolio-based assessments offer a robust means of preserving the authenticity of online assessments in the face of AI chatbots. They align with a comprehensive view of student learning, where the focus is not only on what students know, but also on how they think, learn, and apply knowledge (Ifelebuegu, 2023). Our research shows that interest in portfolios is growing, with 22% of educators considering this method for future use, which reflects a recognition of this method's value.

A promising approach for addressing AI-assisted writing is using **tracking systems**, such as Google Docs and other collaborative platforms, which allow teachers to monitor the development of a student's essay in real time. These systems capture every keystroke, revision, and edit, creating a detailed record of the writing process. Educators can observe student processes, making it easier to determine if the work is genuinely their own or if external content, possibly generated by an AI tool, has been pasted into the document. However, while tracking systems can be highly effective, students can still use AI tools outside of the tracked document to generate content. Additionally, these systems require students to consistently work within a digital framework, which may not always align with their natural writing habits or preferred methods of brainstorming and drafting.

Currently, 14% of educators report using tracking systems like Google Docs to monitor students' writing processes and detect possible AI involvement. Furthermore, interest in these systems is growing, with 28% of respondents considering their adoption for future use. This trend reflects a recognition of the value of tracking

systems in encouraging genuine effort and oversight of students' development. However, English teachers' opinions on the effectiveness of these systems remain mixed. While 14% rate tracking systems as "Very Effective," 24% are neutral. Many teachers appreciate these systems' transparency, though some question their ability to detect AI use fully. With a mean effectiveness rating of 3.76, tracking systems are considered moderately useful but may not completely resolve concerns about AI-assisted writing.

While tracking systems like Google Docs can support academic integrity and encourage authentic student effort, they may not entirely address the complexities posed by AI-assisted writing. As Screti (2023) suggests, balancing technical monitoring with thoughtful feedback and flexibility is essential to foster a more authentic and stress-free learning environment.

Integrating AI tools into English teaching strategies has emerged as an innovative approach to enhancing student engagement and writing proficiency. This strategy allows students to use **AI-assisted applications for tasks like editing and content generation**, providing a controlled environment where they can experiment with technology while developing their writing skills. In our survey, 28% of educators rated AI-assisted editing assignments as "Very Effective," with 48% rating them as "Somewhat Effective." This high level of approval, reflected in a mean score of 4.04 and a low standard deviation of 0.72, highlights the effectiveness of using AI as a supportive tool that does not overshadow students' contributions. A smaller but significant portion (16%) of teachers currently assign tasks where students use AI to generate content, followed by further editing, a method that is projected to grow to 28% in future applications. This approach aligns with research findings that AI tools like QuillBot and other collaborative AI applications can enhance students' writing by offering suggestions and structure, without compromising creativity or logical reasoning (Amyatun & Kholis, 2023; Fang et al., 2023; Fauziah & Minarti, 2023). These findings emphasize a balanced approach, where AI becomes a resource to aid learning, improve skills, and increase confidence while maintaining students' active role in the writing process.

The most effective way to assess AI-influenced writing may be through a **hybrid approach** that combines various assessment methods to achieve a more comprehensive evaluation. For instance, integrating AI detection software with oral defenses provides multiple layers of verification. The software can flag potentially AI-generated content, while the oral defense serves as a secondary check, allowing educators to confirm the student's understanding and authorship. Hybrid approaches offer the flexibility to adapt to various teaching environments and student needs. They allow educators to combine traditional and digital assessments while acknowledging the role of AI in education. However, these methods require considerable teacher involvement and institutional support. Educators need training in AI detection tools and supplementary methods like oral defenses for hybrid strategies to be effective.

Ultimately, educators strongly support hybrid approaches. 52% of respondents rated hybrid methods as "Very Effective," making them the most reliable strategy for addressing AI-related challenges. With the highest mean score of 4.42 and a low standard deviation of 0.67, hybrid methods are regarded as the best solution for assessing student essays in an AI-enhanced academic landscape.

Innovative assessment strategies are essential in addressing the challenges posed by AI-enhanced writing. Whether through tracking systems, oral defenses, AI-detection software, hybrid methods, or collaborative writing tasks, educators are finding new ways to maintain academic integrity in a world where AI tools are increasingly accessible. Each method offers unique strengths and limitations, and the most effective approach will likely involve a combination of these strategies to ensure that student assessments remain fair, accurate, and reflective of their true abilities. As AI

technology advances, assessment methods must also adapt to keep education focused on authenticity, integrity, and real learning.

Conclusions

In the age of AI, English classes face unique challenges and opportunities that require educators to adapt their assessment tools and teaching approaches thoughtfully. Based on our research results, we may conclude that English teachers point out the benefits of AI tools for enhancing certain aspects of student writing, such as grammar and proofreading, while also expressing concerns about overreliance on these tools and the potential impact on academic integrity. The findings reveal that teachers' familiarity with AI tools is relatively high, and most have encountered or suspected AI-influenced student work. However, assessing the authenticity of AI-assisted essays remains a significant challenge, leading many educators to use balanced approaches in AI use, adapted to student level and task type.

Among the assessment methods explored, hybrid approaches, which combine traditional and AI-specific techniques, emerged as the most effective strategy for evaluating AI-influenced essays. Teachers approve methods like in-class handwritten essays, oral defenses, and originality-focused criteria as they promote authenticity and reduce the likelihood of AI misuse. Respondents increasingly use AI detection software, though concerns persist over its reliability and accuracy. Tracking systems and portfolio assessments are viewed as supportive methods that could potentially enhance oversight of the writing process but are not seen as comprehensive solutions.

Implications and Recommendations

Based on our research and other studies (Batista et al., 2024; Farrelly & Baker, 2023; Moorhouse et al., 2023; Attipoe, 2024), we offer the following implications and recommendations for English educators navigating AI-enhanced learning environments.

Focusing on critical thinking and creativity. AI tools like chatbots and language models can help students generate text quickly, which is a big challenge for teachers regarding fair assessment. To ensure students are actively engaged in the English classroom, teachers should focus on assessments that evaluate critical thinking, creativity, and the writing process rather than only assessing final products. AI can be used as a brainstorming tool, but students should analyze and refine ideas independently.

Adopting hybrid assessment models that combine traditional methods with AI detection and tracking systems. These approaches, together with promoting AI literacy and ethical use, can help maintain academic integrity while using the benefits of AI tools to enhance student learning and writing development.

Implementing process-based assessments. This approach emphasizes the cognitive steps involved in writing, such as planning, evaluating information, and revising essays. For example, students could be required to submit drafts, outline their brainstorming activities, and reflect on revisions. Process-focused assessments can help teachers measure students' growth, thinking, and engagement.

Encouraging self-reflection. Self-reflection assignments where students document how they used AI tools and how they arrived at their final work can promote transparency and accountability. Teachers can evaluate students on the quality of their final submission and the reflection and growth evident in their process.

Ethical considerations and responsibility in AI use. Educators are responsible for teaching students to use AI ethically, emphasizing that AI should enhance rather than replace their own ideas. Clear guidelines on ethical AI usage should be developed and implemented in the curriculum.

Limitations

The sample size of 50 participants from one institution may restrict the generalizability of the findings, limiting their applicability to broader educational settings beyond this context. Additionally, due to the rapid evolution of generative AI tools, these findings may become outdated quickly, necessitating future studies to address ongoing advancements in AI.

Perspectives for Future Work

Future research should expand the sample size and explore more diverse educational settings to ensure the findings' broader applicability. Additionally, longitudinal studies that examine how educators' perceptions of assessment methods evolve as AI tools become more sophisticated and prevalent in classrooms are needed.

Bibliographic references

- Alexander, K., Savvidou, C., & Alexander, C. (2023). Who wrote this essay? Detecting AI-generated writing in second language education in higher education. *Teaching English with Technology*, 23(2), 25-43. <https://doi.org/10.56297/buka4060/xhld5365>
- Amyatun, R. L., & Kholis, A. (2023). Can Artificial Intelligence (AI) like QuillBot AI Assist Students' Writing Skills? Assisting Learning to Write Texts using AI. *ELE Reviews: English Language Education Reviews*, 3(2), 135-154. <https://doi.org/10.22515/elereviews.v3i2.7533>
- Attipoe, S. G. (2024). Project management pedagogy: Cultivating critical thinking skills in higher education. *Advanced Education*, 12(24), 151-172. <https://doi.org/10.20535/2410-8286.296878>
- Batista, J., Mesquita, A., & Carnaz, G. (2024). Generative AI and higher education: Trends, challenges, and future directions from a systematic literature review. *Information*, 15(11), 676. <https://doi.org/10.3390/info15110676>
- Cimermanova, I., & Jelavic, D. (2024). Exploring the role of portfolios in fostering student course engagement. *Advanced Education*, 12(24), 118-129. <https://doi.org/10.20535/2410-8286.300555>
- Fang, X., Ng, D.T.K., Leung, J.K.L. et al. (2023). A systematic review of artificial intelligence technologies used for story writing. *Educ Inf Technol*, 28, 14361-14397. <https://doi.org/10.1007/s10639-023-11741-5>
- Farrelly, T., & Baker, N. (2023). Generative artificial intelligence: Implications and considerations for higher education practice. *Education Sciences*, 13(11), 1109. <https://doi.org/10.3390/educsci13111109>
- Farrelly, T., & Baker, N. (2023). Generative Artificial Intelligence: Implications and Considerations for Higher Education Practice. *Education Sciences*, 13(11), 1109. <https://doi.org/10.3390/educsci13111109>
- Fauziah, Z., & Minarti, S. (2023). Artificial Intelligence-Assisted Writing Application for Improving Writing Skills of Islamic Education Students. *Jurnal Kependidikan Islam*, 13(2), 119-123. <https://doi.org/10.15642/jkpi.2023.13.2.119-123>
- Fiialka, S., Kornieva, Z., & Honcharuk, T. (2023). ChatGPT in Ukrainian Education: Problems and Prospects. *International Journal of Emerging Technologies in Learning (iJET)*, 18(17), pp. 236-250. <https://doi.org/10.3991/ijet.v18i17.42215>
- Fleckenstein, J., Meyer, J., Jansen, T., Keller, S. D., Köller, O., & Möller, J. (2024). Do teachers spot AI? Evaluating the detectability of AI-generated texts among student essays. *Computers and Education: Artificial Intelligence*, 6, 100209. <https://doi.org/10.1016/j.caeai.2024.100209>
- Fyfe, P. (2023). How to cheat on your final paper: Assigning AI for student writing. *AI & Society*, 38, 1395-1405. <https://doi.org/10.1007/s00146-022-01397-z>
- Guo, K., Pan, M., Li, Y., & Lai, C. (2024). Effects of an AI-supported approach to peer feedback on university EFL students' feedback quality and writing ability. *The*

- Internet and Higher Education*, 63, 100962.
<https://doi.org/10.1016/j.iheduc.2024.100962>
- Ifelebuegu, A. O. (2023). Rethinking online assessment strategies: Authenticity versus AI chatbot intervention. *Journal of Applied Learning & Teaching*, 6(2). Victoria University Kampala, Uganda. <http://journals.sfu.ca/jalt/index.php/jalt/index>
- Kim, N.J & Kim, M.K. (2022). Teachers' Perceptions of Using an Artificial Intelligence-Based Educational Tool for Scientific Writing. *Frontiers in Education*, 7, 755914. <https://doi.org/10.3389/educ.2022.755914>
- Kizilcec, R. F., Huber, E., Papanastasiou, E. C., Cram, A., Makridis, C. A., Smolansky, A., Zeivots, S., & Radulescu, C. (2024). Perceived impact of generative AI on assessments: Comparing educator and student perspectives in Australia, Cyprus, and the United States. *Computers and Education: Artificial Intelligence*, 7, 100269. <https://doi.org/10.1016/j.caeai.2024.100269>
- Krishna, K., Song, Y., Karpinska, M., Wieting, J., & Iyyer, M. (2023). Paraphrasing evades detectors of AI-generated text, but retrieval is an effective defense. *ArXiv*, abs/2303.13408. <https://doi.org/10.48550/arXiv.2303.13408>
- Lukianenko, V. V., Shastko, I. M., & Korbut, O. H. (2024). Evaluating AI Detection Tools for Academic Integrity in Higher Education. *Scientific Innovations and Advanced Technologies*, 5(33), 970-978. [https://doi.org/10.52058/2786-5274-2024-5\(33\)-970-978](https://doi.org/10.52058/2786-5274-2024-5(33)-970-978)
- Malik, A. R., Pratiwi, Y., Andajani, K., Numertayasa, I. W., Suharti, S., Darwis, A., & Marzuki. (2023). Exploring artificial intelligence in academic essay: Higher education student's perspective. *International Journal of Educational Research Open*, 5, 100296. <https://doi.org/10.1016/j.ijedro.2023.100296>
- Matsuda, P.K., Jeffery, J.V. (2012). Voice in Student Essays. In: Hyland, K., Guinda, C.S. (eds), *Stance and Voice in Written Academic Genres*. Palgrave Macmillan, London. https://doi.org/10.1057/9781137030825_10
- Moorhouse, B. L., Yeo, M. A., & Wan, Y. (2023). Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education: Artificial Intelligence*, 4, 100151. <https://doi.org/10.1016/j.caeo.2023.100151>
- Plate, D., Melick, E., Hutson, J., & Edele, S. (2024). *Generative AI in the English composition classroom: Practical and adaptable*. Routledge.
- Putra, M. (2023). AI writing correction tools: Teachers and students' perception. *Jurnal Pemikiran dan Penelitian Pendidikan*, 21(1), 5. <http://journal.uinmataram.ac.id/index.php/tatsqif>
- Santra, P., & Majhi, D. (2023). Scholarly Communication and Machine-Generated Text: Is it Finally AI vs AI in Plagiarism Detection?. *SRELS Journal of Information Management*. <https://doi.org/10.17821/srels/2023/v60i3/171028>
- Screti, F. (2023). Using Google Docs for guided Academic Writing assessments: students' perspectives. *Language Learning in Higher Education*, 13, 153 - 174. <https://doi.org/10.1515/cercles-2023-2007>
- Sharples, M. (2022). Automated essay writing: An AIED opinion. *International Journal of Artificial Intelligence in Education*, 32(4), 1119–1126. <https://doi.org/10.1007/s40593-022-00300-7>
- Smerdon, D. (2024). AI in essay-based assessment: Student adoption, usage, and performance. *Computers and Education: Artificial Intelligence*, 7, 100288. <https://doi.org/10.1016/j.caeai.2024.100288>
- Soohwan, L., & Song, K. S. (2024). Teachers' and students' perceptions of AI-generated concept explanations: Implications for integrating generative AI in computer science education. *Computers and Education: Artificial Intelligence*, 7, 100283. <https://doi.org/10.1016/j.caeai.2024.100283>

- Sullivan, M., Kelly, A. & McLaughlan, P., (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning & Teaching*, 6(1), 1-10. <https://doi.org/10.37074/jalt.2023.6.1.17>
- Sweeney, S. (2023). Who wrote this? Essay mills and assessment – Considerations regarding contract cheating and AI in higher education. *The International Journal of Management Education*, 21(2), 100818. <https://doi.org/10.1016/j.ijme.2023.100818>
- Tossell, C. C., Tenhundfeld, N. L., Momen, A., Cooley, K., & de Visser, E. J. (2024). Student perceptions of ChatGPT use in a college essay assignment: Implications for learning, grading, and trust in artificial intelligence. *IEEE Transactions on Learning Technologies*, 17, 1069-1081. <https://doi.org/10.1109/TLT.2024.3355015>
- Waltzer, T., Pilegard, C., & Heyman, G. D. (2024). Can you spot the bot? Identifying AI-generated writing in college essays. *International Journal for Educational Integrity*, 20(11). <https://doi.org/10.1007/s40979-024-00158-3>
- Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., et al. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity*, 19(26). <https://doi.org/10.1007/s40979-023-00146-z>

Words: 7605

Characters: 54 261 (30 standard pages)

Assoc. prof. Valentyna Lukianenko
Department of English
Engineering at the National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute.”
Ukraine
lukianenko.valentyna@lil.kpi.ua
ORCID: 0000-0003-3748-2616)

prof. Zoia Kornieva
Department of Theory
Practice and Translation of English,
Dean of the Faculty of Linguistics at the
National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute.”
Kornieva.Zoia@lil.kpi.ua
Ukraine
ORCID: 0000-0002-8848-4323