Educators' and Students' Perceptions of the Use of Writing Surveillance Tools for Academic Integrity and Writing Assessment in Palestinian Higher Education: Draftback as a Case

Montaser Motia Ujvari[1] & Alia Odeh[2,*]

Accepted Manuscript, In press

Abstract: Objectives: This study investigates the integration of Draftback, a browser-based writing surveillance tool, into writing assessment practices in Palestinian universities, with a focus on maintaining academic integrity and enhancing writing assessment amid increased use of generative AI. The research aims to explore students' and educators' perceptions of writing surveillance tools and its potential for identifying plagiarism and enhancing pedagogical support in writing courses. This study represents an early empirical effort to explore the implementation of writing process tracking tools such as Draftback within Palestinian higher education, providing valuable insights into the intersection of surveillance technologies with assessment, teaching practices, and ethical concerns in a context shaped by distinctive political and educational constraints. **Methods:** Mixed methods were employed, including questionnaires completed by 107 English language students and 22 educators from An-Najah National University and Palestine Technical University-Kadoorie, along with analysis of writing scenarios derived from students' Google Docs revision histories. **Results:** Quantitative findings reveal a significant divergence in views: educators expressed more concern than students regarding the ethical risks of AI, and more strongly supported regulation and surveillance to ensure fairness. Both groups, however, acknowledged the developmental value of tracking tools. Qualitative analysis of writing behaviors indicated that patterns such as sudden text appearance and superficial edits were perceived by educators as indicative of AI use. **Conclusions:** The findings suggest that Draftback offers promise for transparent and formative assessment but also raise ethical concerns regarding privacy and punitive applications. In conclusion, a balanced, transparent approach is necessary to foster trust and promote academic integrity while supporting student learning in increasingly digital educational environments.

Keywords: Writing assessment, academic integrity, plagiarism, generative AIs, Writing surveillance tools.

تصورات المعلمين والطلبة حول استخدام أدوات مراقبة الكتابة لضمان النزاهة الأكاديمية وتقييم الكتابة في التعليم العالي الفلسطيني: برنامج"درافت باك" كدراسة حالة

منتصر مطيع أويڤاري[1]، وعالية عودة[2,*]

الملخص: الأهداف: تهدف هذه الدراسة إلى استكشاف دمج أداة "درافت باك "(Draftback)، وهي أداة مراقبة للكتابة تُستخدم عبر المتصفح، في ممارسات تقييم الكتابة في الجامعات الفلسطينية، مع التركيز على الحفاظ على النزاهة الأكاديمية وتعزيز تقييم الكتابة في ظل الاستخدام المتزايد للذكاء الاصطناعي التوليدي، تسعى الدراسة إلى استكشاف تصورات الطلاب والمعلمين حول أدوات مراقبة الكتابة، وإمكاناتها في الكشف عن الانتحال وتعزيز الدعم التربوي في مساقات الكتابة، تمثل هذه الدراسة جهدًا تجريبيًا مبكرًا لاستكشاف تطبيق أدوات تتبع عملية الكتابة مثل "درافت باك" في التعليم العالي الفلسطيني، وتقدّم رؤى مهمة حول تقاطع تقنيات المراقبة مع التقييم والممارسات التعليمية والقضايا الأخلاقية، في سياق يتسم بظروف سياسية وتعليمية فريدة. **المنهجية:** استخدمت الدراسة منهجية مختلطة شملت استبيانات أجريت مع 107 طالبًا في تخصص اللغة الإنجليزية و22 معلمًا من جامعتي النجاح الوطنية وفلسطين التقنية ــ خضوري، إلى جانب تحليل سيناريوهات كتابة مستخلصة من سجلات تعديل مستندات Google الخاصة بالطلاب. **النتائج:** كشفت النتائج الكمية عن تباين ملحوظ في وجهات النظر؛ حيث أبدى المعلمون قلقًا أكبر من الطلاب بشأن المخاطر الأخلاقية للذكاء الاصطناعي، وكانوا أكثر دعمًا لتنظيمه ومراقبته لضمان العدل في التقييم، ومع ذلك، أقرّ كلا الطرفين بالقيمة التطويرية لأدوات التتبع، وأظهرت التحليلات النوعية لأنماط الكتابة أن المعلمين اعتبروا ظهور النصوص فجأة أو وجود تعديلات سطحية مؤشرات على استخدام الذكاء الاصطناعي. **الخلاصة:** تشير النتائج إلى أن أداة "درافت باك" تحمل وعودًا لتقييم شفاف، لكنها تثير أيضًا مخاوف أخلاقية تتعلق بالخصوصية والاستخدام العقابي، وبناءً عليه، توصي الدراسة باتباع نهج متوازن وشفاف لتعزيز الثقة والحفاظ على النزاهة الأكاديمية، مع دعم تعلم الطلاب في بيئات تعليمية رقمية بشكل متزايد.

الكلمات المفتاحية: تقييم الكتابة، النزاهة الأكاديمية، الانتحال، الذكاء الاصطناعي التوليدي، أدوات مراقبة الكتابة.

1 Department of English Language, Faculty of Arts and Educational Sciences, Palestine Technical University-Kadoorie, Tulkarm, Palestine. montaser.owda@ptuk.edu.ps

2 Languages Center, Faculty of Humanities and Educational Sciences, An-Najah National University
Nablus, Palestine
* Corresponding author email: a.odeh@najah.edu

1 قسم اللغة الإنجليزية، كلية الآداب والعلوم التربوية، جامعة فلسطين التقنية - خضوري، طولكرم، فلسطين. montaser.owda@ptuk.edu.ps

2 مركز اللغات، كلية العلوم الانسانية والتربوية، جامعة النجاح الوطنية، نابلس، فلسطين
* الباحث المراسل: a.odeh@najah.edu

Introduction

Writing is a fundamental language skill that plays a crucial role in all aspects of life, especially in academia, where the majority of grades for assignments, exams, and projects are based on writing. Thus, a great deal of attention is given to writing and composition courses at universities and colleges to prepare students for the labor market. After all, a significant amount of work is conducted through writing, including emails, reports, and proposals.

Traditionally, students complete their assignments and exams in class, where invigilators can monitor them to prevent cheating and plagiarism. This is particularly important in general writing courses and language proficiency tests (e.g., IELTS), where the focus is on writing conventions and structure. However, there are circumstances in which writing assignments and exams are conducted at home. This was particularly true during the COVID-19 pandemic (Romaniuk and Łukasiewicz-Wieleba, 2021).

For Palestinians, this situation has persisted even after COVID due to several challenges facing the education sector in Palestine. These challenges are manifested in the closure of university campuses across the West Bank as a result of constant Israeli raids and the excessive installation of checkpoints that prevent both students and teachers from moving to and from their schools and universities (Scott et al., 2024; Scott et al., 2025; Ujvari, 2022). They are also evident in the destruction of most universities in the Gaza Strip, in an event that has been described as 'educide' (Iriqat et al., 2025), referring to the intentional targeting and dismantling of educational infrastructures in war-torn areas (Alousi, 2022). Therefore, testing and assessment have to be carried out at home where educators cannot monitor students' writing process or the resources, they use to complete their assignments and tests.

With the rapid advancement of generative AI technologies, detecting plagiarism in academia has become increasingly difficult, even with plagiarism detection software, which sometimes tend to fail to detect content generated by AI due to its remarkable originality (Khalil and Er, 2023), and it occasionally flags innocent students (Fowler, 2023). Furthermore, it has become significantly more challenging to assess students' actual writing process due to the disconnection between educators and their students. Given these growing concerns about academic integrity in digital learning environments, it is crucial to explore innovative tools that can enhance writing assessment while mitigating plagiarism and unauthorized assistance.

One such tool that enables the real-time playback of students' writing processes is Draftback, a browser extension that "can track, record, and/or quantify students' writing processes and render them as both reports and "replays" (Fernandes and McIntyre, 2025, p. 1). This tool offers educators valuable insights into students' revision patterns, writing progress, and possible instances of academic dishonesty. However, integrating such tools in testing and assessment raises critical questions about their effectiveness, pedagogical significance, and ethical considerations. This study seeks to examine how writing surveillance tools such as Draftback can be utilized to improve assessment in Palestinian universities, especially in writing courses, while maintaining academic integrity. Specifically, the research will explore the following central question:

1. How can writing surveillance tools enhance the assessment of student writing and support academic integrity, from the perspective of both educators and students?

2. How do Palestinian English language educators and students perceive the use of AI in writing assignments?

3. What writing patterns do English language educators in Palestinian universities consider indicative of potential AI use in assignments based on monitoring students' writing processes?

Literature Review

AI in Education: Benefits and Limitations in Language Learning and Teaching

The rise of generative AIs such as ChatGPT has drastically changed the face of academia. It has introduced numerous benefits, compounded by several risks. This is especially true for undergraduate students who seek assistance in writing their assignments, whether a paragraph, an essay, a report, or a project. Nevertheless, the feasibility of ChatGPT as a content-generating tool is sometimes debated. Marcus and Davis (2020) describe ChatGPT as a "fluent spouter of bullshit." Bender et al. (2021) describe ChatGPT as a *stochastic parrot* that constructs language sequences by predicting word patterns based on extensive training data, yet lacks genuine comprehension of the meaning behind the text. Although these claims might have been true at the time, it is undeniable that the ability of generative AIs to produce human-like content is improving every year. Moreover, AIs such as ChatGPT still have the merit of being "fluent" and "parrot-like," which is precisely what students yearn for when writing their assignments for language courses, especially composition ones. This makes the assessment of assignments and exams a challenge for instructors, whose job is to evaluate student learning and understanding.

Generative AIs are said to have numerous advantages in language learning and teaching. They can enhance instruction by fostering interactive and adaptable learning experiences (Kohnke et al., 2023), as they can "identify the meaning of a word in context, correct and explain language mistakes, create texts in various genres (e.g., emails, stories, recipes), develop quizzes, annotate texts, and offer dictionary definitions, example sentences, and translations" (p. 539).Generative AIs can also assist teachers in developing lesson plans (van der Berg & du Plessis, 2023), especially those in the early stages of their careers, boosting their confidence and providing valuable experience in instructional design. Cooper (2003) adds that AI "is likely to be a useful tool for educators designing science units, rubrics, and quizzes" (p. 444). Barber et al. (2021) note that ChatGPT can support assessment by generating personalized exams and quizzes tailored to individual students' needs and abilities. Similarly, Bocul and Sangkawong (2024) highlight that ChatGPT "displays substantial promise as an [Automated Writing Evaluation] tool, offering distinct features such as a human-like interface, consistency, efficiency, and scalability" (p. 1).

However, Manning et al. (2025) note that ChatGPT's grading of written assessments can differ from human raters, as it tends to be more lenient and gives higher scores for essays, making human assessment irreplaceable, at least for now. ChatGPT can also be a valuable tool for enhancing English grammar learning (Annamalai and Bervell, 2025) by identifying grammatical, spelling, and punctuation errors, offering students immediate feedback on their writing (AlAfnan et al., 2023). Some students may find this particularly beneficial, as it provides real-time constructive suggestions, personalized feedback and support, and a self-directed learning experience (Bsharat et al., 2025; Rasul, 2023; Gao, 2021), especially when instructors are not readily available for assistance. Another benefit of generative AIs like ChatGPT is their ability to enhance students' problem-solving and soft skills, including key 21st-century competencies (Assaft et al., 2024).

Despite their undeniable benefits in language learning and teaching, generative AIs pose serious risks to the academic process, including excessive dependence by students, ethical concerns about data privacy, and the reduction of human agency (Mahjoubi et al., 2025; Abu Elnasr et al. 2025). Rowe

(2004) argues that students consistently take advantage of online assessments through plagiarism, which gives them an edge over their peers and results in higher marks. Generative AIs such as ChatGPT have introduced students to advanced technologies with significant potential, yet they primarily facilitate academic dishonesty (Qasem, 2023); these tools allow students to generate automated essays, instantly solve complex problems, and evade conventional plagiarism detection systems more effectively. For instance, issues of plagiarism were detected just two months after ChatGPT's release, resulting in excessive reliance on this tool and leading academics to describe it as a "plague on education" (Weissman, 2023).

For educators, assignment assessment has become a moral dilemma, as teachers must determine whether an assignment genuinely reflects the student's understanding and effort or if it was created by AI, in order to ensure fairness in assessment among their students (Kiryakova and Angelova, 2023). This process can be more time-consuming. Therefore, educators are increasingly burdened with the challenge of detecting AI-generated content, which has become more difficult due to advances in technology.

Academic Integrity and Plagiarism

Academic integrity is a fundamental pillar of the academic profession. It encompasses the values, behaviors, and ethical standards expected of faculty in their roles as teachers and researchers. As higher education systems expand globally, academic integrity has become increasingly significant, not only to preserve institutional credibility but also to promote a professional culture that actively resists misconduct, plagiarism, and unethical behavior. Integrity, rooted in the Latin term *integritas*, implies moral wholeness, blending honesty, accountability, and respect for others (MacIntyre, 1981; Fjellstrom, 2005).

The significance of academic integrity also extends to the reputation of academic institutions. Universities now compete on an international stage where incidents of academic misconduct can undermine trust in higher education systems (Altbach, 2004). Therefore, the rise of global university rankings has intensified the importance of plagiarism detection, making academic integrity not only an issue of personal and professional ethics but also one of strategic importance. Consequently, scholars and policymakers have called for clearer codes of conduct, faculty development initiatives, and institutional structures to uphold ethical standards in teaching and research (Braxton et al., 2011). As Macfarlane et al. (2014) suggest, cultivating academic integrity is essential for sustainable educational growth and must be reinforced by both ethical reflection and structural reform.

Cotton et al. (2024) note that mitigating students' use of AI in completing their assignments could be carried out through educating students on the risks of plagiarism, requiring students to submit a draft of their work for review before the final submission, using plagiarism detection tools, and closely monitoring students' work. However, in the Palestinian context, implementing some of these methods could be challenging. Some Palestinian undergraduate students are unaware of the risks of plagiarism (Hassan, 2024). Hamamra et al. (2024) note that some Palestinian university students, especially the less motivated ones, tend to rely on submitting texts generated by AI without reading or proofreading them, which can suppress their creativity and critical thinking skills. They also note that this has led some educators to change their methods of assessment by focusing on oral presentations, in-class participation, and exams. Although these methods could be useful and ensure that students cannot access generative AI tools during assessment, the complete shift to distance education, including lectures, assignments, and exams, poses a greater challenge.

Traditional plagiarism detection tools, such as Turnitin and other freely available detection tools, have been widely used to detect plagiarism and the use of AI. However, these commonly used tools in academic settings have been found to be neither precise nor reliable, exhibiting low accuracy rates as they tend to classify text as human-written rather than effectively identifying AI-generated content (Weber-Wulff, 2023; Pegoraro et al., 2023; van Oijen, 2023). Therefore, relying solely on AI detection tools could be problematic and may result in potentially unfair assessments of students' assignments.

The present study

This study aims to investigate how writing surveillance tools, namely, Draftback, can be integrated into writing assessment practices in Palestinian higher education to uphold academic integrity amid the growing use of generative AI tools like ChatGPT. It seeks to explore the perceptions of both students and educators regarding the ethical, pedagogical, and practical implications of using such tools. Additionally, the study examines whether observable writing patterns, as captured through writing revision histories, can be indicative of AI-generated content. By addressing these objectives, the research aims to contribute to a more transparent and formative approach to writing assessment in digital learning environments, particularly in the unique and challenging context of Palestine.

METHODS

A convenience sampling method was employed to recruit participants from two Palestinian universities: An-Najah National University and Palestine Technical University-Kadoorie. Questionnaires designed by the researchers were distributed to 107 university students aged between 18 and 23, nine of whom were males, all majoring in English Language and Literature at two aforementioned Palestinian universities. The unbalanced number of female and male participants in this study is expected, as Palestinian females tend to significantly outnumber males in the humanities and social sciences, according to the Palestine Central Bureau of Statistics (2024). Most participants were from the governorates of Nablus, Tulkarm, and Jenin.

The questions explored students' attitudes toward the use of surveillance software designed to track their writing process. Among the participants, 98 reported that they had never received formal training or workshops on AI, 4 reported that they had, and 5 were unsure. Additionally, when asked whether their teachers discuss the use of AI in assignments, responses varied: 48 students reported that their lecturers always discuss the use of AI, 40 said that their teachers sometimes do, and 19 stated that their lecturers do not discuss AI use in assignments at all.

The eleven Likert-scale questions required students to indicate their level of agreement or disagreement with each statement, with 5 representing "strongly agree" and 1 representing "strongly disagree." The statements, along with their mean and standard deviation values, are reported in Tables 1, 2, 3 and 4.

Furthermore, seven samples of students' essays from an Academic Writing course taught at Palestine Technical University-Kadoorie were collected and analyzed using Draftback. Students were instructed to submit their essays in the form of Google Docs and were explicitly asked to type the entire essay rather than copying and pasting it. They were not informed that their writing process would be observed without explicitly mentioning the name of the software. The students consented on using samples of their essays for research purposes while ensuring anonymity. After analyzing the students' writing processes, several scenarios were developed based on the findings.

Another questionnaire, consisting of eleven Likert-scale questions, was distributed to 22 English language lecturers (11 males and 11 females) teaching English language courses at An-Najah National

University and Palestine Technical University–Kadoorie. The participants were evenly divided between holders of MA and PhD degrees. The questionnaire investigated the educators' perceptions of the potential role of a virtual writing surveillance tool that tracks the entire writing process and records every edit and change over time to improve student performance and combat plagiarism.

The questionnaire also included an additional section presenting scenarios developed from students' actual writing. In this section, lecturers were asked to assess the likelihood of plagiarism in the scenarios using the following categories and their corresponding numerical values: "Highly Likely (4)," "Moderately Likely (3)," "Unlikely (2)," and "No Evidence of AI Use (1)." Teachers were then required to explain their assessments in one sentence. The mean and standard deviation values for their responses are reported in the tables below. When asked whether they had received formal training on the use of AI in teaching, only six participants reported that they had, while the remaining participants indicated they had not.

To assess the internal consistency of the questionnaire items, Cronbach's alpha was calculated using R Studio. The overall reliability coefficient for the scale was found to be 0.79, indicating good internal consistency and suggesting that the items measured a coherent underlying construct.

The inclusion of both students and educators was essential to capture the dual perspectives that shape writing assessment practices, especially when integrating surveillance tools like Draftback, and to allow a comparative analysis of how writing surveillance is perceived from both pedagogical and experiential perspectives. Prior studies show that students and educators often differ in how they perceive assessment tools, which can impact both implementation and effectiveness (Carless, 2006; Nicol & Macfarlane-Dick, 2006).

RESULTS

Quantitative Analysis: University students vs educators' perceptions about the use of AI and Writing Surveillance tools in writing assessment

A series of Mann-Whitney U tests were conducted to examine differences between students (N = 107) and educators (N = 22) in their responses to 14 Likert-scale items regarding the use of AI tools in academic writing. Significant differences were found on several items. Educators were significantly more likely than students to believe that AI tools negatively impact learning outcomes (U = 778, p < .01), make it easier to plagiarize or paraphrase without understanding (U = 790, p < .01), and may reduce students' writing skills over time (U = 786, p < .01). Educators were also more concerned that AI tools hinder critical thinking (U = 860, p < .05), considered AI use unethical (U = 761, p < .01), and viewed it as a form of plagiarism (U = 571, p < .01). Moreover, educators placed significantly more importance on institutional policy regarding AI usage (U = 622, p < .01) and supported regulation of AI tools to ensure fair assessment (U = 750, p < .01). Conversely, students were significantly more open to allowing AI tools in academic work compared to educators (U = 1647, p < .01).

While students and educators expressed differing views on several aspects of AI in academic writing, there were also meaningful areas of agreement. Both groups generally believed that AI tools can support students' writing development, particularly by improving the quality of writing and reducing writing anxiety through feedback and suggestions. They also showed similar levels of support for the idea that AI has the potential to enhance academic integrity when used to assist with research and writing tasks. In addition, students and educators shared the view that tracking students' writing processes can be beneficial for assessing development.

Table (1): Student vs Teacher's perceptions on using AI in writing. Significant P values are highlighted in bold.

Questionnaire item	Students' response		Educators' response		Statistical analysis	
	M	SD	M	SD	P value	U
1. I believe that AI tools can improve the quality of students' writing in English.	4.	0.7	3.7	1	0.3	1320
2. I believe that the use of AI tools by students in assignments may negatively impact their learning outcomes.	3.2	1.06	3.9	0.9	**<0.01**	778
3. I believe that AI tools make it easier for students to paraphrase or copy content without fully understanding it	3.7	1.04	4.4	0.6	**<0.01**	790
4. I believe that AI tools should be allowed in academic assignments.	3.3	1	2.9	1.2	0.12	1412
5. I am concerned that excessive reliance on AI tools may reduce students' writing skills over time.	3.9	1	4.5	0.6	**<0.01**	786
6. I believe that AI tools can help reduce students' writing anxiety by providing suggestions and feedback.	4.1	0.7	3.8	0.7	0.11	1399
7. I am concerned that students' use of AI tools may hinder the development of essential academic skills, such as critical thinking and writing.	3.5	1	4.1	0.8	**<0.05**	860
8. I believe that using AI tools in assignments could be considered unethical.	3.1	1.1	3.9	1	**<0.01**	761
9. I believe that using AI tools in academic tasks could be considered a form of plagiarism and cheating.	2.8	0.9	3.9	1	**<0.01**	571
10. I believe it is important for my institution to have clear policies regarding the use of AI tools in academic assignments.	3.7	1	4.5	0.8	**<0.01**	622
11. The use of AI tools should be regulated to ensure fair assessment.	3.9	0.9	4.5	0.7	**<0.01**	750
12. I believe that the use of AI tools may make it difficult for me to assess students' true understanding of the material.	3.7	1.1	4.1	0.8	0.10	930
13. I believe that educators should be more open to allowing students to use AI tools in academic work.	3.6	0.9	2.9	0.9	**<0.01**	1647
14. I believe that AI tools have the potential to enhance academic integrity by assisting students with research and writing.	3.7	0.8	3.4	1	0.18	1371

The study also explored attitudes toward writing tracking and surveillance tools through responses to six Likert-scale questions. Educators were significantly more supportive than students of using writing tracking tools to improve assessment transparency (U = 754, p < .01), identify areas where

students need more practice (U = 856, p < .05), and increase confidence in tracking student progress (U = 833, p < .05). A significant difference was also found regarding attitudes toward the use of such tools for punitive or surveillance purposes, with educators expressing greater acceptance than students (U = 579, p < .01). Interestingly, neither group strongly perceived the use of writing tracking tools as an invasion of privacy (U = 1453, p = .70), suggesting a shared sense of acceptance or neutrality when these tools are used for academic purposes. Additionally, general support for writing tracking as a developmental tool showed no statistically significant difference between groups.

These shared perceptions highlight a foundation of mutual understanding between students and educators, particularly in recognizing the supportive role AI can play in the writing process, as well as a measured stance on ethical concerns related to digital surveillance in educational contexts. Overall, the results indicate that while both groups recognize the potential benefits of AI, educators express significantly more concern about its negative impact and are more supportive of regulation and writing surveillance mechanisms than students.

Table (2): Student vs Teacher's perceptions on using writing surveillance tools. Significant P values are highlighted in bold.

Questionnaire item	Students' response		Educators' response		Statistical analysis	
	M	**SD**	**M**	**SD**	**P value**	**W**
15. Tracking students' writing processes from start to finish would be a useful tool for assessing their development in English writing.	4	0.8	4.4	0.7	0.06	901
16. believe that writing tracking tools can enhance transparency in the assessment process and reduce opportunities for academic dishonesty.	3.7	0.8	4.3	0.7	**<0.01**	754
17. I believe that tracking students' writing processes can help identify areas where students need more practice or improvement.	3.9	0.9	4.3	9.3	**<0.05**	856
18. It is acceptable if writing tracking tools were used for purposes other than academic improvement (such as punitive measures or surveillance against cheating).	2.8	0.8	3.8	0.9	**<0.01**	579
19. I believe that using tools to track students' writing could be an invasion of their privacy.	3.2	1.1	2.7	1.2	0.7	1453
20. I would feel more confident in (my/students') progress if I could track and analyze their writing history.	3.4	1	3.9	0.9	**<0.05**	833

Furthermore, the questionnaires for educators and students included items tailored to each group's context. Responses from university students generally indicated a positive outlook toward the use of

writing tracking tools. However, one item reflecting student discomfort with instructors monitoring their writing process and revision history received a neutral average score (M = 3.0, SD = 1.0), suggesting a mild level of unease.

In contrast, students responded more favorably to items highlighting the potential advantages of tracking tools. They strongly supported the idea that such tools could enhance their awareness of their own writing processes (M = 3.9, SD = 0.9) and encourage greater accuracy in their work (M = 3.8, SD = 0.9). The notion that teacher access to revision history might boost motivation received moderate agreement (M = 3.4, SD = 1.0).

Transparency was a key concern for students, with the highest level of agreement with the importance of clear communication about how instructors would use the data collected through these tools (M = 3.9, SD = 0.8). Additionally, students somewhat agreed that writing tracking could help bridge the gap between their self-evaluations and instructors' assessments (M = 3.5, SD = 0.8).

Table (3): Mean and standard deviation values for the questionnaire items tailored specifically to university students about the use of writing surveillance tools.

Questionnaire item	M	SD
1. I would feel uncomfortable if my teachers used tools to track the writing process and the revision history of my assignments.	3	1
2. I believe that if teachers have the ability to track my writing, it will help me become more aware of my writing process.	3.9	0.9
3. I would feel more motivated to improve my writing if my teachers had access to my writing process and revision history.	3.4	1
4. I feel that using writing tracking tools will encourage me to work on my tasks more accurately	3.8	0.9
5. I believe that writing tracking tools should be used transparently, with clarification on how teachers will use the data.	3.9	0.8
6. I believe that writing tracking tools can help bridge the gap between students' self-assessment and teachers' grading	3.5	0.8

Palestinian university lecturers demonstrated strong support for incorporating writing tracking tools into academic practice. The highest level of agreement was linked to the belief that such tools enable instructors to provide more effective support to students struggling with specific aspects of writing (M = 4.4, SD = 0.7). Educators also strongly agreed that tracking students' writing processes is beneficial for monitoring their development over time (M = 4.3, SD = 0.7) and for identifying individual strengths and weaknesses in writing (M = 4.3, SD = 0.9).

In addition, lecturers recognized the value of gaining insights into students' writing behaviors. They agreed that tracking tools could help uncover patterns such as tendencies to delay or focus effort and distinguish between superficial edits and meaningful revisions (M = 4.2 for both items; SD = 0.7 and 0.8, respectively). There was also broad agreement that access to students' writing processes could enable instructors to provide more targeted and constructive feedback (M = 4.2, SD = 0.8).

Table (4): Mean and standard deviation values for the questionnaire items tailored specifically to university educators about the use of writing surveillance tools.

Questionnaire item	M	SD
Tracking students' writing process from beginning to end can be a useful tool for monitoring the development of my writing over time.	4.3	0.7
I believe that access to students' writing processes would help me better understand the development of their writing and guide my feedback.	4.2	0.8

I believe that tracking students' writing processes will help them identify their strengths and weaknesses in English writing.	4.3	0.9
I believe that writing tracking tools can help me identify patterns in students' writing habits (e.g., when they tend to procrastinate or focus more).	4.2	0.7
I believe that tracking writing processes will help determine whether students are engaging in genuine revisions or they are just making superficial changes.	4.2	0.8
I believe that tracking writing processes will allow instructors to provide better support to students who may be struggling with certain aspects of writing.	4.4	0.7

Qualitative Analysis: Detection of AI-Generated Writing by Tracking the Writing Process

The scenario-based evaluation of AI-generated writing provided insights into how writing tracking tools can help identify potential instances of AI-assisted work. Educators were presented with the following scenarios and asked to judge the likelihood that AI was used to generate each one.

Scenario 1: Sudden Appearance of Text A

In this scenario, a student submitted an essay. When reviewing their writing history through the tracking tool, it was observed that the entire text appeared instantly, rather than being typed gradually, indicating a copy-paste procedure. The English language educators found the appearance of text instantly with no evidence of typing it and without carrying out any revisions is indicative of plagiarism and use of AI (M = 3.5, SD = 0.6). Teacher #17 elaborated:

"The sudden appearance of the whole text suggests that the student did not type the text gradually. This could be evidence that the student copied and pasted the entire text without even reading it."

Scenario 2: Large Sections Pasted in Multiple Segments

In this scenario, it was noticed that large sections of the text were pasted in multiple segments with minimal revisions. Only a few words were changed afterward. Similarly to scenario 1, English language educators also deemed the text in this scenario to be generated by AI (M = 3.5, SD = 0.6). Teacher #1 explained this scenario by stating that the text is likely to be generated by AI, but the students tried to make few changes to the text to avoid detection. Similarly, Teacher #12 elaborated that the student was merely selective of what they had copied. However, Teacher #17 made the following statement regarding this scenario:

"If large sections of text are pasted in multiple segments with minimal revision, it suggests that the student may have copied the content from external sources. However, this does not prove that the text was generated using an AI tool."

Scenario 3: Typing Followed by Deletion and Pasting

In this scenario, the student initially typed a few sentences but then deleted them and pasted a significantly longer passage. The pasted text remained mostly unchanged throughout the writing process. The educators participating in this study reported a plausible likelihood of use of AI in generating the content of the assignment (M= 3.2, SD = 0.9). Teachers #12 and #2 predicted that the student in this scenario may have lost motivation to continue writing and then took an easier approach and used AI to complete the assignment. According to Teacher #8:

"It seems that the student was not using AI at the beginning. However, later on, it appears he was struggling or felt frustrated which made him use AI to complete the task."

Scenario 4: Consistent, Gradual Typing with Few Pauses

In this scenario, the student's essay was written gradually over time, with regular pauses and revisions. Initially, there were spelling mistakes, but some of them were later corrected. The text still

had some grammatical and punctuation errors which were not corrected. The educators decided that the chances of using AI were low (M = 2, SD = 0.9). Still, some of them replied to the likelihood of using AI to generate content in the written assignment as "Moderately likely". Teacher #3 suggests that there is a chance that the student made their own revisions while Teachers #1 and #8 suggest that writing, deleting, and editing some parts is normal in the writing process, and that correction of spelling is a feature integrated in MS word, which does not prove the use of AI.

Scenario 5: Deliberate Induced Spelling Mistakes

In this scenario, it was observed that the text appeared instantly, with no evidence of typing. Then, some words were erased and were replaced with incorrect spelling after they had been written correctly. The replacements of correctly written words with incorrectly written ones were consistent and appeared to be done intentionally rather than by accident. The educators evaluating the likelihood of using AI in writing the assignment showed strong agreement that the text is AI generated (M= 3.7, SD = 0.7). Several respondents made the same remarks: the student in this scenario tried to deceive the teacher through inducing some spelling mistakes to make them think that the text is human-generated. Teacher #22 remarks that:

"The student tried to deceive the teacher to make them think the text is the student's authentic writing by showing that their text includes some spelling mistakes, which is something AI can never do without being instructed to."

Scenario 6: Sudden Shift in Writing Style

In this scenario, the student's writing was consistent throughout the majority of the writing process, but suddenly there was a significant shift in tone, vocabulary, and sentence structure. This abrupt change was detected in the middle of the essay and persisted until the end. Respondents described the likelihood of using AI in the assignment as "Moderately likely" (M = 3.1, SD = 0.7). Teacher #3 commented that AI was used when the shift in style took place. However, other teachers claim that this cannot be a clear indication of use of AI. Teacher #1 demonstrates that a change of writing style is not enough evidence to prove that the student used AI. Teacher #7 further demonstrates that it could depend on whether the writing process happened at one time since the writing style could change after going back to writing after a period of time.

Scenario 7: Long Time Between Initial Writing and Final Submission

In this scenario, the student began writing the essay several days before the deadline but, after drafting an initial version, left the document mostly untouched for a long period. Then, just hours before the deadline, a large amount of text—mostly typed—was added and revised in one session. The results from educators were mixed. While some teachers believed there was no evidence of AI use, others reported a high likelihood of AI assistance. Teacher #11 suggested that writing large segments at once after a period of inactivity could indicate AI help in completing the assignment. However, other teachers, namely #1, #8, and #17, thought the student was simply a procrastinator or a "last-minute player" who started motivated but rushed the assignment as the deadline approached.

The scenario-based evaluation of AI-generated writing demonstrates how writing tracking tools can help educators identify potential instances of AI assistance by analyzing students' writing behaviors. Cases such as the sudden appearance of text or large sections being pasted with minimal revision (Scenarios 1 and 2) were strongly associated with AI use, as they suggest copying from an external source without a natural writing progression. Similarly, instances of deleting initial writing attempts and replacing them with pasted text (Scenario 3) raised concerns about AI use, as students may resort

11

to AI when struggling to complete assignments. Educators also noted that deliberate spelling errors (Scenario 5) could serve as a deceptive strategy to mask AI use. More complex scenarios, such as shifts in writing style (Scenario 6) or last-minute bulk writing (Scenario 7), generated mixed responses, highlighting the limitations of tracking tools in definitively proving AI involvement. These results also reveal disparities in how educators judge plagiarism attempts, with some appearing more rigid than others. Therefore, while these tools provide valuable insights into the writing process, their effectiveness in detecting plagiarism attempts cannot always be guaranteed and may result in false positives.

DISCUSSION

The results collected from both students and educators on the use of tracking tools in the assessment of written assignments highlight the potential benefits and challenges associated with implementing writing tracking tools such as Draftback in educational settings. While students recognize the advantages of these tools for self-assessment and skill development, their awareness of how tracking can aid in identifying areas for improvement and increasing accountability is tempered by concerns about its use in plagiarism detection that may lead to disciplinary actions. In contrast, Palestinian educators generally view writing tracking tools positively, emphasizing their role in monitoring student progress, identifying writing habits, and ensuring transparency in assessment. These findings are consistent with Khlaif et al. (2024), who found that university teachers in the Middle East viewed generative AI tools as both promising and problematic, expressing strong concern about ethical misuse, diminished critical thinking, and the need for clear institutional guidelines on AI integration. Fernandes and McIntyre (2025) note writing surveillance tools such as Draftback can be beneficial to writers to establish the authenticity of their writing, where teachers find it as a means to mitigate reliance on AI in their courses.

However, educators' lower concern for privacy issues and more favorable attitudes toward punitive measures may promote a teacher-student relationship characterized by adversarial perspectives on plagiarism. Furthermore, reliance on writing surveillance tools could encourage a narrow definition of what writing should look like and may lead to increased anxiety, distrust, and fear among students (Fernandes & McIntyre, 2025; Giray et al., 2025).

Recent research has raised concerns about the increasing reliance on educational technologies for monitoring and assessment, which may signal a shift toward surveillance capitalism in academia (Zuboff, 2019; Chu, 2024). In this context, writing surveillance tools like Draftback do more than merely track writing behaviors, they participate in the commodification of student data under the guise of pedagogical improvement. While these tools offer valuable insights that may support formative assessment, they also risk normalizing surveillance practices that prioritize institutional control over student agency. This risk is particularly acute in vulnerable contexts like Palestine, where education is already entangled with issues of freedom and occupation. Embedding surveillance into writing assessment in such settings could reinforce asymmetrical power relations and deepen students' feelings of insecurity. As Hamamra, Qabaha, and Daragmeh (2022) argue, the panoptic nature of digital learning environments transformed education into a mechanism of control, where both instructors and students internalized the gaze of an "invisible Other." This dynamic undermined the emancipatory potential of higher education by replacing pedagogical care with fear-driven self-discipline. Furthermore, the suppression of free speech in Palestinian universities, fueled by political surveillance, internal factionalism, and institutional complicity, has created a climate of fear and self-censorship that

undermines democratic discourse and critical pedagogy, making the call for parrhesiastic, truth-telling speech all the more urgent (Hamamra and Gould 2024). Therefore, it is crucial that educators and institutions critically interrogate not only the effectiveness but also the broader implications of writing surveillance tools in a digital education landscape.

CONCLUSION

In conclusion, writing tracking tools offer both benefits and challenges in writing assessment. While they support transparency and help mitigate plagiarism in online education, which has become the norm in Palestinian universities, concerns over privacy and punitive applications create tensions between students and educators. The scenario-based evaluation highlights how these tools can help identify AI-assisted writing but also reveals their limitations, including the risk of false positives. Ultimately, a balanced approach that prioritizes transparency, clear guidelines, and ethical implementation is essential for ensuring their effective and fair use in education, and assessment guidelines should be established when using such tools.

This study has some limitations. First, the educators in this study did not have full access to students' writing processes, which led to the development of scenarios. They were only presented with a summary of the writing process and had to assess the likelihood of plagiarism based on these scenarios. In addition, future research could benefit from increasing the number of participating English language educators to provide a more comprehensive perspective. Another limitation involves the validation of the questionnaire. Although the questionnaire developed by the researchers was grounded in the literature review, it was not formally validated by experts. Nonetheless, the Cronbach's alpha coefficient indicated that the questionnaire is reliable.

Draftback, as a tool, also has its own limitations, the most significant being that it functions as a browser extension. This means that tracking students' writing processes can only be done through Google Docs, which may conflict with institutions that require students to submit assignments through a designated Learning Management System (LMS). Future research should explore strategies to balance pedagogical benefits with ethical considerations, ensuring that writing tracking tools enhance education rather than impose control.

It is recommended that educational institutions prioritize transparency by clearly communicating the purpose and scope of these tools to students, including whether they could be used as surveillance tools for imposing punitive measures. If these tools are used to mitigate plagiarism and detect copying from AI, clear guidelines should be implemented. This aligns with Khlaif et al.'s (2024) recommendation that institutional policies be developed to ensure AI tools are used ethically and constructively, supporting both student development and academic fairness.

Disclosure statement
- **Ethical Approval and Consent to Participate:** Not applicable
- **Availability of Data and Materials:** Data are available upon request from the corresponding author .
- **Authors Contribution:** All authors contributed to the design, data analysis ,
- **Conflict of Interest:** There are no conflicts of interest
- **Funding:** No funding .
- **Acknowledgments:** We thank the participating students and educators for their valuable contributions.

Open Access

REFERENCE

- Sobaih Abu Elnasr E.., Mahjoubi, M. N., Brini, R., & Al-Qutaish, A. A. (2025). Transforming Education with AI: An Exploratory Study of Faculty Insights on ChatGPT's Opportunities and Risks. *An-Najah University Journal for Research-B (Humanities)*, *39*(4). DOI: 10.35552/0247.39.4.2370
- AlAfnan, M. A., Dishari, S., Jovic, M., & Lomidze, K. (2023). Chatgpt as an educational tool: Opportunities, challenges, and recommendations for communication, business writing, and composition courses. *Journal of Artificial Intelligence and Technology, 3(2)*, 60-68.
- Alousi, R. (2022). Educide: the genocide of education a case study on the impact of invasion and conflict on education. *The Business and Management Review, 13(2)*, 8-9. https://doi.org/10.24052/bmr/v13nu02/art-28
- Altbach, P. (2004). The question of corruption. *International Higher Education*, (34).
- Annamalai, N., & Bervell, B. (2025). Exploring ChatGPT's role in English grammar learning: A Kolb model perspective. *Innovations in Education and Teaching International*, 1-17.
- Assaf, R., El-Qanni, A., Alsurakji, I. H., & Assad, M. (2024). The ChatGPT Dilemma: Opportunities vs. Risks in Engineering Education. *An-Najah University Journal for Research-B (Humanities)*, *38*(10), 1985-2002. DOI: 10.35552/0247.38.10.2285
- Barber, M., Bird, L., Fleming, J., Titterington-Giles, E., Edwards, E., & Leyland, C. (2021). Gravity assist: Propelling higher education towards a brighter future. *Office for Students*. https://www.officeforstudents.org.uk/publications/gravity-assist-propelling-higher-education-towards-a-brighter-future/
- Bsharat, T. R., Ahmad, I. S., & Khlaif, Z. (2025). Advantages and Disadvantages of Access to ChatGPT Among University Students: A Systematic Literature Review. *Enhancing Learning Experiences With Digital Tools: AI, ChatGPT, and Virtual and Augmented Reality*, 23-46.

– Braxton, J. M., Proper, E. M., & Bayer, A. E. (2011). *Professors behaving badly: Faculty misconduct in graduate education*. JHU Press.

– Bucol, J. L., & Sangkawong, N. (2024). Exploring ChatGPT as a writing assessment tool. *Innovations in Education and Teaching International*, 1-16.

– Carless, D. (2006). *Differing perceptions in the feedback process. Studies in Higher Education, 31*(2), 219–233. https://doi.org/10.1080/03075070600572132

– Chu, N. E. (2024). *The influence of surveillance capitalism on educational technology adoption* (Doctoral dissertation, University of British Columbia).

– Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 61(2)*, 228-239.

– Fernandes, M., & McIntyre, M. (2025). Drafting defensively, documenting authorship: An analysis of Draftback and Grammarly Authorship. *Computers and Composition, 76*, 102926.

– Fjellstrom, R. (2005). Respect for persons, respect for integrity: remarks for the conceptualization of integrity in social ethics. *Medicine, Health Care and Philosophy*, *8*, 231-242.

– Fowler, G. A. (2023). We tested a new ChatGPT-detector for teachers. It flagged an innocent student. *The Washington Post, 3.*

– Gao, J. (2021). Exploring the feedback quality of an automated writing evaluation system Pigai. *International Journal of Emerging Technologies in Learning (iJET, 16(11)*, 322–330. https://doi.org/10.3991/ijet.v16i11.19657

– Giray, L., Sevnarayan, K., & Ranjbaran Madiseh, F. (2025). Beyond Policing: AI Writing Detection Tools, Trust, Academic Integrity, and Their Implications for College Writing. *Internet Reference Services Quarterly*, *29*(1), 83-116.

– Hamamra, B., Qabaha, A., & Daragmeh, A. (2022). *Online education and surveillance during COVID-19 pandemic in Palestinian universities*. *International Studies in Sociology of Education*, 31(4), 446–466. https://doi.org/10.1080/09620214.2021.2016473

– Hamamra, B., Mayaleh, A., & Khlaif, Z. N. (2024). Between tech and text: The use of generative AI in Palestinian universities—a ChatGPT case study. *Cogent Education, 11(1)*, 2380622.

– Hamamra, B., & Gould, R. R. (2024). *Free speech and democracy in Palestinian universities: A call for parrhesiastic speech*. *Educational Philosophy and Theory*, 56(13), 1317–1331. https://doi.org/10.1080/00131857.2024.2391857

– Hassan, A. (2024). Knowledge, attitude, opinion, perspective, and agreement of Palestinian medical students on strategies/recommendations to curb plagiarism: A multicenter cross-sectional study. *Accountability in Research, 31(8)*, 1085-1106.

– Iriqat, D., Alousi, R., Aldahdouh, T. Z., AlDahdouh, A., Dankar, I., Alburai, D., ... & Hassoun, A. (2025). Educide amid conflict: The struggle of the Palestinian education system. *Quality Education for All, 2(1)*, 1-19.

– Khalil, M., & Er, E. (2023, June). Will ChatGPT get you caught? Rethinking plagiarism detection. *International Conference on Human-Computer Interaction* (pp. 475-487). Cham: Springer Nature Switzerland.

- Khlaif, Z. N., Ayyoub, A., Hamamra, B., Bensalem, E., Mitwally, M. A., Ayyoub, A., & Shadid, F. (2024). *University teachers' views on the adoption and integration of generative AI tools for student assessment in higher education*. *Education Sciences*, 14(10), 1090. https://doi.org/10.3390/educsci14101090

- Kiryakova, G., & Angelova, N. (2023). ChatGPT—A challenging tool for university professors in their teaching practice. *Education Sciences, 13(10)*, 1056.

- Macfarlane, B., Zhang, J., & Pun, A. (2014). Academic integrity: a review of the literature. *Studies in higher education*, *39*(2), 339-358.

- MacIntyre, A. 1981. After virtue. London: Duckworth.

- Manning, J., Baldwin, J., & Powell, N. (2025). Human versus machine: The effectiveness of ChatGPT in automated essay scoring. *Innovations in Education and Teaching International*, 1-14.

- Nicol, D. J., & Macfarlane-Dick, D. (2006). *Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. Studies in Higher Education,31(2), 199– 218.* doi:10.1080/03075070600572090

- Pegoraro, A., Kumari, K., Fereidooni, H., & Sadeghi, A. R. (2023). To ChatGPT, or not to ChatGPT: That is the question! *arXiv.* https://doi.org/10.48550/arXiv.2304.01487

- Qasem, F. (2023). ChatGPT in scientific and academic research: Future fears and reassurances. *Library Hi Tech News, 40(3)*, 30-32. https://doi.org/10.1108/lhtn-03-2023-0043

- Rasul, T., Nair, S., Kalendra, D., Robin, M., de Oliveira Santini, F., Ladeira, W. J., ... & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning and Teaching, 6(1)*, 41-56.

- Romaniuk, M. W., & Łukasiewicz-Wieleba, J. (2021). Challenges of administering university examinations remotely during the COVID-19 pandemic. *E-Mentor. Czasopismo Naukowe Szkoły Głównej Handlowej w Warszawie*, *90*(3), 22-31.

- Scott, H., Ujvari, M. M., & Smith, M. (2024). "Our identity is our dignity": Digital transformations: Palestinian aspirations, idealism, reality, and pragmatism. *Education as Change, 28(1)*, 1-14.

- Scott, H., Ujvari, M., Bakeer, A. M. A., & Shanaa, K. (2025). Sumud as connected learning: Towards a collective digital commons in Palestine. *Journal of Interactive Media in Education, 2025(1)*, 8, pp. 1–15. https://doi.org/10.5334/jime.877

- Ujvari, M. M. (2022). Linguistic Landscape in the West Bank: Road Signs as Manifestations of Occupation. *International Journal of Language and Literary Studies*, *4*(1), 374-387. https://doi.org/10.36892/ijlls.v4i1.881

- van Oijen, V. (2023). AI-generated text detectors: Do they work? *SURF Communities.* https://communities.surf.nl/en/ai-in-education/article/ai-generated-text-detectors-do-they-work.

- Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., & Waddington, L. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity, 19(1)*, 1-39.

- Weissman, J. (2023, February 9). ChatGPT is a plague upon education. *Inside Higher Ed.* https://www.insidehighered.com/views/2023/02/09/chatgpt-plague-upon-education-opinion

- Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power, edn. *PublicAffairs, New York*.