**FORUM DE L'ENSEIGNANT Journal**
e-ISSN: 2437-0576 ISSN: 1112-5101
**Volume: 22, Special Issue: 1 / October 2026 pp. 82-99**

# When the Pen Is Not Entirely Ours: Ethical Considerations and Academic Integrity in Technology-Supported Writing

**Received: 04/07/2026; Accepted: 27/10/2026; Published: 31/10/2026**

Houda FADEL [1],*

*1 Mohamed Lamine Debaghine Setif 2 University (Algeria), h.fadel@univ-setif2.dz.*

ORCID: (https://0009-0007-2950-6805.org/)

**Abstract:**
This paper aims to examine the ethical considerations and academic integrity challenges raised by technology-supported writing, with particular attention to generative artificial intelligence (AI) in higher education and English as a Foreign Language (EFL) context. Adopting a qualitative, narrative literature review methodology based on a targeted search of the scholarly literature, the study synthesizes recent scholarship on academic integrity, technology-supported writing, and second-language authorship, together with emerging research from Algerian higher education. It addresses four central questions: how generative AI reshapes notions of authorship and originality; to what extent transparency rather than detection can serve as the primary ethical criterion for evaluating AI-assisted writing; how reliable and equitable AI-detection tools are, particularly for EFL and other multilingual writers; and which institutional and pedagogical responses best safeguard academic integrity while preserving opportunities for technology-supported language learning. Drawing on this synthesis, the paper proposes practical principles, guidelines, and illustrative scenarios for the ethical use of writing technologies by students, educators, and institutions. It concludes that academic integrity in the age of AI-supported writing depends less on prohibition or detection than on transparency, clearly articulated expectations, and assessment designs that make honest engagement with one's own learning the most reasonable choice available to students.

**Keywords:**
Academic integrity;
Technology-supported writing;
Generative AI;
Plagiarism;
EFL writing;

* *Corresponding author: Houda FADEL*
*Doi:*

## 1. Introduction

Writing has never been a purely unmediated act. Long before the arrival of generative artificial intelligence, writers relied on dictionaries, style guides, spellcheckers, and, later, grammar-checking software to shape their prose. Word processors themselves were once viewed with suspicion by some educators, who worried that the ease of cutting, pasting, and revising text electronically would erode students' capacity for sustained, effortful composition. Each new writing technology, from the typewriter to the spellchecker to predictive text, has prompted a similar cycle of concern, adaptation, and eventual normalization (Godwin-Jones, 2022). What distinguishes the current moment is not the mere presence of technology in the writing process, but the degree to which contemporary tools can generate, rather than simply correct or support, substantial portions of a text.

The emergence of large language models capable of producing fluent, coherent, and often highly convincing academic prose in response to a brief prompt has unsettled long-standing assumptions about what it means to write, to know, and to demonstrate learning through text (Perkins, 2023). Unlike earlier writing technologies, which primarily assisted a human author in refining their own ideas, generative AI tools can produce original-seeming content with minimal human input, raising the question of where legitimate assistance ends and academic misconduct begins. This question is not merely technical; it is deeply ethical, touching on the values of honesty, trust, fairness, respect, and responsibility that the International Center for Academic Integrity (2021) identifies as foundational to academic communities.

This paper examines the ethical considerations and academic integrity implications of technology-supported writing, with particular attention to the challenges posed by generative AI in higher education and EFL contexts. It situates current debates within the longer history of writing technology and plagiarism scholarship, examines four interrelated ethical dimensions of AI-supported writing, and considers how institutions and educators might respond in ways that protect academic integrity without foreclosing the legitimate pedagogical benefits that writing technologies can offer. Throughout, the paper argues that ethical clarity, rather than technological prohibition or detection alone, offers the most sustainable path forward for academic writing in a technology-saturated environment. Because EFL learning in Algerian universities is the context that motivates this inquiry, the paper also draws on emerging local research to examine how far international findings travel.

More specifically, this study pursues three aims:

(1) to trace how technology-supported writing has evolved from early automated feedback tools to contemporary generative AI, situating this evolution within

existing plagiarism and academic integrity scholarship (Howard, 1995; Pecorari, 2003; Godwin-Jones, 2022);

(2) to identify and examine the ethical dimensions raised by generative AI in academic writing; and

(3) to derive practical, evidence-based recommendations, including illustrative scenarios, for students, educators, and institutions.

These aims translate into four guiding research questions:

**RQ1:** How does the emergence of generative AI reshape long-standing conceptions of authorship and originality in academic writing?

**RQ2:** To what extent can transparency and disclosure, rather than prohibition or detection, function as the primary ethical criterion for evaluating AI-assisted writing?

**RQ3:** How reliable and equitable are current AI-detection technologies, particularly with respect to EFL and other multilingual writers?

**RQ4:** What institutional and pedagogical responses are best suited to safeguarding academic integrity while preserving legitimate opportunities for technology-supported language learning?

Building on the scholarship reviewed below, the paper advances three guiding propositions. Because a narrative literature review synthesizes and interprets existing scholarship rather than collecting new data, these propositions are arguments to be examined against the reviewed literature, not hypotheses to be tested or confirmed:

**P1:** generative AI blurs, rather than eliminates, the boundary between legitimate writing support and academic misconduct, such that authorship is better understood as a continuum than a binary.

**P2:** transparency-based policies are more ethically defensible and practically sustainable than detection-based enforcement (Eaton, 2023).

**P3:** AI-detection tools can disproportionately disadvantage non-native English and EFL writers, such that their uncritical use risks compounding existing inequities rather than resolving them (Liang et al., 2023).

## 2. Review Methodology

This paper is a narrative literature review. Unlike a systematic review, it does not aim to retrieve every study on the topic; it aims instead to synthesize and critically interpret a purposively selected body of scholarship in order to build a

conceptual argument about authorship, transparency, equity, and detection in technology-supported writing. Sources were identified between [2003] and [ 2025] through searches of Google Scholar, Scopus, ERIC, and the Algerian Scientific Journal Platform (ASJP), supplemented by backward citation tracing from key articles and by searches of the websites of professional and institutional bodies (for example, the International Center for Academic Integrity, APA Style, and Turnitin) for guidance documents and policy statements. Search terms were combined with Boolean operators and included "academic integrity", "plagiarism", "patchwriting", "generative AI", "ChatGPT", "AI detection", "automated writing evaluation", "second language writing", "EFL", "non-native", and "Algeria".

The review covers publications from 2003 to 2025. The starting date reflects the foundational work on patchwriting and second-language authorship that predates generative AI (Howard, 1992, 1995; Pecorari, 2003), whereas most of the remaining sources were published after late 2022, when generative AI tools became widely available in higher education. Sources were included if they (a) addressed academic integrity, writing technologies, AI-text detection, or second-language writing in higher education; (b) were peer-reviewed publications or authoritative institutional, professional, or industry documents; and (c) were available in English. Sources were excluded if they were opinion pieces without argument or evidence, if they did not concern writing or assessment, or if their content could not be checked against the original publication or institutional page. Within these limits, sources were selected for their relevance to the four research questions, their influence on the scholarly debate, their recency, and the strength of their evidence, with empirical studies and primary documents preferred over secondary commentary. Because the aim is conceptual synthesis rather than exhaustive coverage, the review rests on a deliberately focused core of studies, a choice whose consequences are discussed in Section 7.

The selected sources were read in full and analyzed thematically, with themes organized around the four research questions (authorship and originality, transparency and disclosure, detection and equity, and institutional and pedagogical responses). Findings from empirical studies were interpreted with attention to sample size, context, and the tools examined, and the resulting arguments are presented as propositions (P1–P3) rather than as tested hypotheses.

## 3. Literature Review

### a. Academic Integrity: Foundations and Evolving Definitions

Academic integrity has traditionally been understood as a commitment to a set of core values that sustain trust within scholarly communities. The International Center for Academic Integrity (2021) identifies six such values:

honesty, trust, fairness, respect, responsibility, and courage. These values are not abstract ideals; they translate into concrete expectations about how students represent their own work, how educators evaluate that work fairly, and how institutions create conditions in which honest effort is recognized and rewarded. Historic conceptions of academic misconduct centered on relatively straightforward acts such as copying another person's words without attribution, fabricating data, or submitting work completed by someone else on the student's behalf.

The rise of digital writing technologies has complicated this relatively stable picture. As Eaton (2023) argues in her concept of postplagiarism, the growing normalization of AI and other advanced technologies in everyday life means that academic integrity can no longer be adequately understood through inherited, technology-blind definitions of plagiarism. Rather than treating the presence of technology in the writing process as inherently suspect, Eaton (2023) suggests that integrity in a postplagiarism era depends on transparency about how a text was produced and on creating learning environments in which shortcuts offer little genuine advantage. This reframing does not dissolve the concept of academic integrity; rather, it relocates the ethical burden toward disclosure, institutional clarity, and pedagogical design, rather than toward an ever-expanding list of prohibited tools.

**b. Technology-Supported Writing: From Spellcheckers to Generative AI**

The category of "technology-supported writing" spans a wide continuum of tools with very different implications for authorship and academic integrity. At one end are automated writing evaluation (AWE) systems and automated written corrective feedback (AWCF) tools, such as Grammarly, which supply real-time suggestions about grammar, spelling, and style but leave the substantive content and argument to the human writer (Godwin-Jones, 2022). These tools have generally been treated as acceptable, even encouraged, forms of writing support, particularly for second-language writers who benefit from immediate, targeted feedback on surface-level errors.

At the other end of the continuum are generative AI systems, such as ChatGPT, capable of producing extended, coherent text in response to a short prompt, with little to no further input from the human user (OpenAI, 2023). Godwin-Jones (2022) notes that this newer generation of tools represents a qualitatively different kind of writing assistance: rather than refining a text that the writer has already produced, these systems can generate the text itself, raising fundamental questions about whose ideas, whose language, and whose intellectual effort the final product actually represents. Between these two poles lie a range of intermediate tools, including AI-powered paraphrasing software,

machine translation, and hybrid systems that combine generative capability with more traditional editing functions, each of which raises its own distinct ethical questions.

### c. Plagiarism, Patchwriting, and Authorship in the Digital Era

Long before generative AI entered classrooms, applied linguists had already documented the complex and often unintentional ways in which writers, particularly second-language writers, could produce text that resembled plagiarism without deceptive intent. The concept of patchwriting originates in the work of Howard (1992, 1995), who used the term for the practice of copying from a source text and then altering it through deletions, grammatical changes, or one-for-one synonym substitutions, and who argued that this practice is better approached as a pedagogical matter than as a purely juridical one (Howard, 1995). Pecorari's (2003) influential study of postgraduate second-language writers applied and developed the concept for second-language writing, showing that patchwriting can be a common developmental stage in which writers, still building fluency and disciplinary knowledge, copy source material closely while substituting some words and phrases, often as a strategy for managing unfamiliar content and language rather than as intentional deception. This body of research established an important precedent that is highly relevant to the current debate over generative AI: textual similarity to a source, or textual features associated with non-independent authorship, do not necessarily indicate dishonest intent, and blanket punitive responses risk penalizing legitimate language-learning processes.

This precedent matters directly for how the field interprets AI-assisted writing. Just as patchwriting research shifted some scholarly and institutional attention from moralistic judgment toward pedagogical explanation, the emergence of generative AI invites a similar, though not identical, reconsideration of what constitutes authorship. Unlike patchwriting, which typically involves a source text that the writer has read and imperfectly transformed, AI-generated text often has no single retrievable source; it is a statistical synthesis of patterns learned from vast quantities of text. This distinction matters ethically, because it means that AI-generated content cannot be evaluated using the same textual-similarity logic that underlies traditional plagiarism detection, a point that has significant implications for the detection technologies discussed in Section 4.d.

## 4. Ethical Considerations in Technology-Supported Writing

### a. Authorship and Originality

The most fundamental ethical question raised by generative AI in academic writing concerns authorship: who, or what, can meaningfully be said to have written a given text. When a student submits an essay substantially generated by an AI tool as though it were their own independent work, the act raises concerns

not only about honesty but about the underlying purpose of academic writing assignments, which are typically designed to develop and demonstrate the writer's own thinking, research, and command of language (Perkins, 2023). At the same time, few educators would consider it dishonest for a student to use a spellchecker, consult a dictionary, or ask a peer to review a draft for clarity. The ethical difficulty lies in locating where legitimate support ends and substitution of the student's own intellectual effort begins, a boundary that is considerably harder to draw for generative AI than for earlier writing technologies precisely because these tools can operate anywhere along the continuum from minor editorial assistance to wholesale text generation.

**b. Transparency and Disclosure**

A growing consensus in the academic integrity literature holds that transparency, rather than prohibition, should be the primary ethical criterion governing the use of AI in academic writing. The American Psychological Association's APA Style guidance on generative AI reflects this principle at the level of scholarly publishing: it advises authors to describe how they used tools such as ChatGPT and to credit them through an in-text citation and a reference-list entry (McAdoo, 2023), and its more recent guidance encourages fuller disclosure and citation of AI use, noting that transparency is part of sound scholarly reporting (McAdoo et al., 2025), rather than treating AI assistance as undisclosed background support. Applied to student writing, this principle suggests that the ethical status of a given use of AI depends substantially on whether it was used openly, in line with stated course or institutional expectations, or concealed in a way that misrepresents the origin of the work. Eaton's (2023) postplagiarism framework similarly emphasizes that ethical practice in a technology- saturated writing environment depends on explicit, transparent accounting of how a text was produced, rather than on an assumption that any technological assistance beyond spelling and grammar correction is inherently suspect.

**c. Equity and Access**

Ethical analysis of technology-supported writing must also attend to questions of equity. Writing technologies are not equally accessible to all students: premium AI tools, high-quality internet access, and personal devices remain unevenly distributed both within and across countries, meaning that policies assuming universal access to particular tools risk disadvantaging students with fewer resources (Godwin-Jones, 2022). A related equity concern arises from the opposite direction: where AI writing tools are broadly accessible, they may offer particular linguistic advantages to learners writing in a language other than their

first, potentially levelling a playing field that has historically disadvantaged multilingual and EFL writers in English-medium academic contexts. Research on second-language writing supports the view that such tools offer real affordances to these writers, while also pointing to contradictions in their use for language learning (Warschauer et al., 2023). Determining whether and how such tools should be permitted therefore requires balancing concerns about independent skill development against the recognition that language proficiency itself is unevenly distributed for reasons unrelated to a student's effort or ability.

**d. Detection Tools and Their Limits**

Institutions have increasingly turned to AI-detection software as a technological response to the challenges generative AI poses for academic integrity (Perkins et al., 2024b). However, the ethical case for relying heavily on such tools is considerably weaker than their widespread adoption might suggest. Liang et al. (2023) demonstrated that several widely used GPT detectors misclassified more than half of TOEFL essays written by non-native English speakers as AI-generated, while showing near-perfect accuracy on essays written by native English-speaking students, a pattern the authors attribute to detectors penalizing the comparatively constrained lexical and syntactic variety often characteristic of non-native academic writing. This bias carries serious ethical consequences: a detection tool that disproportionately flags the writing of EFL and other multilingual students risks compounding the very inequities that writing technologies might otherwise help to address, and false accusations of AI use can cause significant harm to falsely accused students regardless of an institution's ultimate finding.

Kofinas et al. (2025) add a further complication, reporting from an exploratory experiment that even experienced markers evaluating authentic, real-world-style assessments struggled to reliably distinguish AI-generated or AI-modified work from independently produced student writing, and that the use of more authentic assessment formats did not, by itself, safeguard academic integrity against sophisticated AI use. Taken together, this body of research suggests that AI-detection tools, and even human judgment applied to sophisticated AI-assisted text, cannot be treated as a reliable or ethically sufficient safeguard against misuse. Chechitelli (2023), writing as Turnitin's Chief Product Officer, acknowledges that the false-positive rate of the company's widely used detection system is not zero and advises instructors to apply professional judgment, knowledge of their students, and the context of the assignment. The ethical implication is significant: institutions that treat detection software as a definitive arbiter of misconduct, rather than as one imperfect source of evidence among several, risk substituting one form of injustice for another.

These findings should nonetheless be read in light of their methodological limits and of how far they can be generalized. The evidence reported by Liang et al. (2023) rests on a small corpus of 91 TOEFL essays collected from a Chinese educational forum, compared with 88 essays by US eighth-grade students, and it concerns seven detectors available at the time of the study. The comparison group was therefore school-age native writers rather than university students writing disciplinary prose, the TOEFL sample represents a single timed-test genre and a single learner population, and detectors released or updated since then may behave differently. The authors' link between misclassification and low linguistic variability suggests a mechanism, but it does not establish how a particular institutional tool will perform on a particular cohort. Kofinas et al. (2025) themselves describe their experiment as exploratory: it involved eight markers and 21 writing samples produced with ChatGPT 3.5 in two UK business schools, so its conclusions concern marker judgment in one disciplinary and institutional setting. Evidence on detectors is also still developing: studies of Turnitin's AI indicator and of six widely used detectors indicate that prompting or editing strategies can reduce detection accuracy and that detector output alone is an unsafe basis for allegations of misconduct (Perkins et al., 2024a, 2024b). These limitations do not overturn the concern about bias, but they qualify it. The most defensible conclusion is that detector outputs are unreliable as sole evidence, especially for EFL writers, and that any tool should be validated locally before it informs high-stakes decisions. No evaluation of detector performance on the writing of Algerian EFL students was identified among the sources reviewed here, so the extent of the problem in that context remains an open empirical question (see Section 5.d).

## 5. Academic Integrity Challenges and Institutional Responses

### a. Policy Responses

Higher education institutions have responded to the challenges of generative AI with a wide range of policy approaches, from outright prohibition to conditional permission to active pedagogical integration. Harvard College's Office of Undergraduate Education, for example, requires faculty to inform students of the generative AI policy governing each course and offers draft statements ranging from maximally restrictive to fully encouraging, with the permissive and mixed versions requiring that any AI use be acknowledged and cited; its guidance also describes reliance on automated detection as inadvisable (Harvard College Office of Undergraduate Education, n.d.). This course-by-course, acknowledgment-based approach reflects the transparency principle discussed in Section 4.b. Other institutions, by contrast, have introduced AI-

detection tools into their assessment and academic-misconduct processes (Perkins et al., 2024a, 2024b), despite the documented limitations of such tools. This policy heterogeneity reflects genuine uncertainty within the field about the appropriate balance between restriction and integration, and it means that students moving between institutions, or even between courses within the same institution, may encounter substantially different expectations about acceptable AI use.

Perkins (2023) argues that a blanket institutional ban on generative AI tools is neither feasible nor enforceable, given both the difficulty of reliable detection and the likelihood that such tools will become increasingly embedded in professional writing and research contexts that students are being prepared to enter. This suggests that institutional policy is most effective, and most ethically defensible, when it is context-sensitive: clearly articulated at the level of individual courses and assignments, rather than imposed as a uniform, one-size-fits-all rule that fails to account for disciplinary differences in what constitutes legitimate use of writing technology.

**b. Pedagogical Responses**

Beyond formal policy, many educators have begun to redesign assessments in ways intended to make the honest, independent development of ideas more valuable, and technological shortcuts less appealing, to students. Approaches include process-oriented assessment that evaluates drafts, reflection, and revision history rather than only a polished final product; in-class or oral components that require students to articulate and defend their arguments in real time; and assignments that explicitly incorporate AI tools as objects of critical analysis rather than treating their use as inherently illegitimate. However, Kofinas et al. (2025) caution that authentic, real-world-style assessments alone do not reliably protect academic integrity against sophisticated AI use, suggesting that pedagogical redesign must be paired with clear communication of expectations and, where appropriate, genuine dialogue with students about why independent effort matters for their own learning, rather than functioning as a purely technical fix.

**c. The Role of EFL and L2 Writers**

The ethical stakes of technology-supported writing are particularly pronounced for EFL and other second-language writers, for whom writing technologies can serve legitimate compensatory functions that native English-speaking peers may not require to the same degree. Godwin-Jones (2022) emphasizes that AI-enabled writing tools, when integrated thoughtfully into instruction, can help language learners develop metalinguistic awareness and become informed, critical consumers of writing technology, rather than either uncritical users or students excluded from tools that could genuinely support their

learning. Pecorari's (2003) earlier findings regarding patchwriting reinforce the importance of distinguishing between EFL writers who rely on technological support as part of a legitimate, ongoing language-learning process and those who use such tools to circumvent that learning process altogether. Institutional and pedagogical responses that fail to make this distinction risk either unfairly restricting legitimate support for multilingual writers or, conversely, allowing meaningful skill development to be bypassed entirely.

**d. Evidence from Algerian Higher Education**

Research on technology-supported writing in Algerian higher education is still emerging and consists largely of small-scale studies, but it points to patterns that mirror the international debate. In a questionnaire study of Algerian EFL students, Cheriguene and Dahbi (2025) found that many students use ChatGPT to enhance vocabulary, organize ideas, and overcome language barriers, while a significant portion remain uncertain about the ethical boundaries of such use: 16.9% neither agreed nor disagreed that submitting unacknowledged AI-generated content constitutes plagiarism, and academic pressure and low proficiency emerged as the main factors behind ethically questionable use. Similarly, in an exploratory study of 56 Master 1 EFL students at Batna 2 University, Boulanouar (2025) reports that most respondents (76.9%) used tools such as ChatGPT, Grammarly, or Microsoft Copilot, mainly for grammar, vocabulary, and idea generation; that opinion was divided on whether using AI for assignments amounts to academic dishonesty; and that students voiced concern about over-reliance. At the level of research writing, Brahmi and Nesba (2025) interviewed thirty Algerian doctoral researchers, who valued AI support for tasks such as literature review, translation, and proofreading but stressed the need for caution about research ethics; the authors recommend detailed policies and training for responsible use.

Taken together, these studies suggest that the central difficulty in Algerian higher education may be less a widespread intention to deceive than uncertainty about where acceptable assistance ends, which is precisely the ambiguity that transparency-based policies and explicit, assignment-level expectations are designed to address. They also reinforce the equity argument of Section 4.c, since students report turning to AI tools to compensate for language barriers, and they underline the need for the guidance developed in Section 6. Two cautions apply. First, these studies rely on self-reported data from small samples in particular institutions, so they illustrate rather than establish national patterns. Second, they focus on students' and researchers' practices and perceptions rather than on the performance of AI-detection tools, which means that the extent of detection bias for Algerian EFL writers has yet to be examined empirically.

## 6. Toward Ethical Frameworks for Technology-Supported Writing

### a. Principles for Responsible Use

Drawing on the preceding analysis, several principles can guide the ethical use of technology-supported writing in academic contexts. First, transparency should function as the primary ethical criterion: disclosed and appropriately cited use of writing technologies is categorically different, ethically, from concealed use presented as entirely independent work (Eaton, 2023; McAdoo, 2023). Second, the appropriate level and kind of technological support should be understood as context-dependent, varying by discipline, assignment purpose, and learner population, rather than governed by a single universal rule. Third, detection technologies should be treated as one imperfect source of evidence among several, never as a sole or definitive basis for findings of misconduct, given their documented and substantial bias against non-native English writers (Liang et al., 2023). Fourth, equity considerations, including differential access to tools and differential linguistic need for technological support, should be built into policy design from the outset rather than addressed only after inequitable outcomes become apparent.

### b. Practical Guidelines for Students and Educators

For students, the clearest ethical path is straightforward disclosure: when in doubt about whether a particular use of technology is permitted, asking the instructor rather than assuming permission or guilt is both an ethically sound and a practically protective strategy. For educators, the guidelines emerging from this analysis include stating AI-use expectations explicitly and in writing for each assignment rather than relying on generic institutional policy; treating any AI-detection signal as a prompt for a supportive conversation rather than an automatic disciplinary trigger, particularly for students who may be multilingual writers; and, where possible, designing assessments that make the process of writing, not only the final product, visible and assessable. For institutions, the analysis suggests that policy clarity, faculty training in the current limitations of detection technology, and ongoing revision of guidelines as the technology continues to evolve are essential components of a genuinely ethical institutional response.

### c. Operationalizing the Guidelines: A Tiered Model and Illustrative Scenarios

The guidelines above can be made more concrete through three instruments that educators can adapt to their own courses. The first is a tiered AI-use statement attached to each assignment. At Level 1 (no generative AI), the task is completed without AI tools, as in timed in-class writing or diagnostic tasks. At Level 2 (limited and disclosed support), students may use AI tools for defined purposes,

such as brainstorming or feedback on grammar and vocabulary, and must disclose the use. At Level 3 (integrated use), the assignment requires students to work with AI output, for example by critiquing or revising it, and to submit the prompts and outputs together with a reflection on what they accepted, rejected, and why. This structure parallels the range of draft course policies offered by Harvard College's Office of Undergraduate Education (n.d.), from restrictive to fully encouraging, but makes the permitted uses explicit at the level of the individual assignment.

The second instrument is a short disclosure statement submitted with every Level 2 or Level 3 assignment, stating which tool was used, for what purpose, which parts of the text were affected, and how the student checked and revised the output. Such statements apply to student work the logic of APA Style guidance, which asks authors to describe how AI was used and recommends that they keep a record of their prompts (McAdoo et al., 2025). The third instrument is process evidence, such as dated drafts, annotated revisions, or a brief conference in which the student explains key choices. Process evidence makes the writing process assessable and gives educators a fair basis for dialogue when questions about authorship arise.

The following scenarios illustrate how the tiers and the principles discussed above might guide decisions in an EFL writing course.

**Scenario 1: An unexpected detection score.** A third-year EFL student submits a literature review whose language is markedly more fluent than her in-class writing, and a detector reports a high probability of AI generation. Because detector output is imperfect and may be biased against non-native writers, the score is treated as a prompt for a conversation, not as proof. The instructor invites the student to discuss her drafting process, asks her to explain key passages and to show her notes and drafts, and takes into account the AI-use tier of the assignment. If AI support was permitted at Level 2 but not disclosed, the first response is educational, clarifying the disclosure requirement; a misconduct procedure is considered only if substantial evidence of misrepresentation emerges.

**Scenario 2: Disclosed language support.** On a Level 2 assignment, a Master's student drafts her own argument and then asks a chatbot to suggest corrections and alternative phrasings, which she reviews and partly adopts. She declares this use in her disclosure statement. The use is consistent with the assignment's expectations and is assessed on the quality of her argument and revisions. The instructor may also use the case to build metalinguistic awareness by asking her to compare the original and revised sentences and to explain which changes she accepted and why (Godwin-Jones, 2022).

**Scenario 3: Undisclosed substitution.** On a Level 1 task, a student submits text generated by a chatbot from the assignment prompt with only minor edits, and cannot explain the content when asked. Because expectations were explicit and the concern does not rest on a detection score alone, the case is addressed under the institution's misconduct procedures, with a proportionate response and an opportunity for the student to explain; where appropriate, support for language and study skills is offered alongside any sanction.

**Scenario 4: Adopting a detector.** A department considers screening all submissions with an AI detector. Before adopting it, the department tests the tool on a sample of student writing produced before generative AI tools became widely available, including work by EFL students, to estimate false-positive rates for its own cohorts; trains staff to interpret scores; states in writing that scores cannot be the sole basis for allegations; and gives students a clear route to respond. If the pilot shows unacceptable error rates for EFL writers, the department prioritizes assessment redesign and dialogue over screening.

## 7. Discussion

The ethical and academic integrity challenges surrounding technology-supported writing cannot be resolved through appeals to a fixed, technology-independent notion of "originality," because writing has always been supported, to varying degrees, by tools external to the writer's unaided mind. What generative AI has changed is the degree to which those tools can substitute for, rather than merely support, the writer's own intellectual labor, and the difficulty of reliably detecting when that substitution has occurred. This shift demands a correspondingly updated ethical vocabulary, one that centers transparency, context-sensitivity, and equity rather than a search for increasingly sophisticated detection technologies that, as the evidence reviewed here suggests, can be unreliable as sole evidence and may be biased against precisely the students, non-native English writers among them, who may most need legitimate technological support (Liang et al., 2023; Kofinas et al., 2025).

The implications extend beyond individual instructor decisions to the broader mission of higher education institutions. If academic integrity is fundamentally about sustaining trust within a scholarly community (International Center for Academic Integrity, 2021), then institutional responses that rely primarily on surveillance and punitive detection risk eroding exactly the trust they are meant to protect, particularly when detection tools risk producing disproportionate false positives for already marginalized student populations. A more sustainable path treats students as partners in defining legitimate use, in line with Eaton's (2023) postplagiarism framework, which locates integrity not in the absence of technology but in the honesty and clarity with which its use is acknowledged.

For EFL contexts specifically, the discussion carries an additional layer of significance. Writing technologies, used transparently and pedagogically, hold genuine potential to support language development rather than undermine it, but this potential can only be realized within institutional and classroom cultures that distinguish clearly between technology as a scaffold for learning and technology as a substitute for it. Achieving this distinction in practice requires sustained dialogue between researchers, educators, and students, rather than a one-time policy pronouncement, given how rapidly the underlying technologies continue to evolve.

Several limitations qualify these arguments and indicate how far they can be applied. As a narrative review based on a purposively selected corpus of English-language sources, the paper is open to selection bias and does not claim exhaustive coverage (Section 2). The empirical evidence on detector bias comes from a small number of studies conducted outside Algeria with tools available in 2023 (Section 4.d), and the Algerian studies discussed in Section 5.d are small-scale and rely on self-reported data. Generative AI and detection technologies also change quickly, so specific findings may date faster than the ethical arguments they inform. For these reasons, the propositions advanced here should be read as arguments supported by the reviewed literature, not as tested conclusions, and their applicability to particular institutions, including Algerian universities, should be examined empirically, as the research agenda in Section 8 proposes.

## 8. Conclusion

This paper has examined the ethical considerations and academic integrity challenges raised by technology-supported writing, situating current debates about generative AI within a longer history of writing technology and plagiarism scholarship. In response to RQ1, the reviewed literature supports P1: generative AI blurs rather than eliminates the boundary between legitimate assistance and misconduct, so that authorship is better understood as a continuum than as a binary. In response to RQ2, it supports P2: transparency-based approaches, grounded in disclosure and clearly articulated expectations, appear more ethically defensible and practically sustainable than prohibition or covert detection. In response to RQ3, it lends support to P3: AI-detection tools have been shown, in studies with important limitations, to misclassify the writing of non-native English writers, so that their uncritical use risks compounding rather than resolving existing inequities. In response to RQ4, the review suggests that institutional and pedagogical responses are most promising when they combine policy clarity, assessment redesign, and explicit attention to the needs of EFL and multilingual writers. Together, these arguments identify authorship, transparency,

equity, and the limits of detection as four interrelated ethical dimensions of the issue.

The distinctive contribution of the paper lies in bringing these dimensions together from the standpoint of EFL writers. It connects the longer scholarship on patchwriting and second-language authorship to current debates on generative AI, treats detection bias as an equity problem rather than a purely technical one, and translates the resulting arguments into a tiered, assignment-level model with illustrative scenarios, informed by emerging evidence from Algerian higher education.

Students are encouraged to disclose their use of writing technologies and to seek clarification when expectations are uncertain. Educators are encouraged to state AI-use expectations for each assignment, to treat detection outputs as one imperfect source of evidence, and to make the writing process visible and assessable. Institutions are encouraged to adopt context-sensitive policies, to train staff in the limits of detection tools, and to build equity considerations into policy design from the outset.

Future research should examine how students and educators in EFL contexts, including Algerian universities, experience and negotiate AI policies in practice; test the performance of AI-detection tools on local EFL student writing; compare policy models across institutional and national contexts; and follow EFL learners' language development longitudinally under different regimes of technology-supported writing. Because the tools will continue to change, these questions will require ongoing reconsideration. What should remain constant is the commitment that academic integrity has always represented: honesty, trust, fairness, respect, responsibility, and the courage to uphold them.

## References

Boulanouar, S. (2025). Integrating artificial intelligence in creative writing: Boon or barrier for EFL students? *Passerelle, 14*(1), 264–273. https://asjp.cerist.dz/en/downArticle/531/14/1/280127

Brahmi, M., & Nesba, A. (2025). Artificial intelligence in academic research: Exploring potential and ethical considerations in Algerian universities. *The Journal of El-Ryssala for Studies and Research in Humanities, 10*(1), 976–985. https://asjp.cerist.dz/en/downArticle/223/10/1/265594

Chechitelli, A. (2023). *Understanding false positives within our AI writing detection capabilities* [Blog post]. Turnitin. https://www.turnitin.com/blog/understanding-false-positives-within-our-ai-writing-detection-capabilities

Cheriguene, A., & Dahbi, A. (2025). Between assistance and abuse: Exploring the ethical use of ChatGPT in Algerian EFL context. *Journal of Legal and Economic Research, 8*(2), 78–111. https://asjp.cerist.dz/en/article/268667

Eaton, S. E. (2023). Postplagiarism: Transdisciplinary ethics and integrity in the age of artificial intelligence and neurotechnology. *International Journal for Educational Integrity, 19*(1), Article 23. https://doi.org/10.1007/s40979-023-00144-1

Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology, 26*(2), 5–24. https://hdl.handle.net/10125/73474

Harvard College Office of Undergraduate Education. (n.d.). *Generative AI guidance*. Harvard University. Retrieved October 29, 2026, from https://oue.fas.harvard.edu/faculty-resources/generative-ai-guidance/

Howard, R. M. (1992). A plagiarism pentimento. *Journal of Teaching Writing, 11*(2), 233–246.

Howard, R. M. (1995). Plagiarisms, authorships, and the academic death penalty. *College English, 57*(7), 788–806. https://doi.org/10.2307/378403

International Center for Academic Integrity. (2021). *The fundamental values of academic integrity* (3rd ed.). https://www.academicintegrity.org/fundamental

Kofinas, A. K., Tsay, C. H.-H., & Pike, D. (2025). The impact of generative AI on academic integrity of authentic assessments within a higher education context. *British Journal of Educational Technology, 56*(6), 2522–2549. https://doi.org/10.1111/bjet.13585

Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns, 4*(7), Article 100779. https://doi.org/10.1016/j.patter.2023.100779

McAdoo, T. (2023, April 7). *How to cite ChatGPT* [Blog post]. APA Style. https://apastyle.apa.org/blog/how-to-cite-chatgpt

McAdoo, T., Denneny, S., & Lee, C. (2025, September 9). *Citing generative AI in APA Style: Part 1—Reference formats* [Blog post]. APA Style. https://apastyle.apa.org/blog/cite-generative-ai-references

OpenAI. (2023). *ChatGPT* (Mar 14 version) [Large language model]. https://chat.openai.com/chat

Pecorari, D. (2003). Good and original: Plagiarism and patchwriting in academic second-language writing. *Journal of Second Language Writing, 12*(4), 317–345. https://doi.org/10.1016/j.jslw.2003.08.004

Perkins, M. (2023). Academic integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching & Learning Practice, 20*(2), Article 7. https://doi.org/10.53761/1.20.02.07

Perkins, M., Roe, J., Postma, D., McGaughran, J., & Hickerson, D. (2024a). Detection of GPT-4 generated text in higher education: Combining academic judgement and software to identify generative AI tool misuse. *Journal of Academic Ethics, 22*(1), 89–113. https://doi.org/10.1007/s10805-023-09492-6

Perkins, M., Roe, J., Vu, B. H., Postma, D., Hickerson, D., McGaughran, J., & Khuat, H. Q. (2024b). Simple techniques to bypass GenAI text detectors: Implications for inclusive education. *International Journal of Educational Technology in Higher Education, 21*(1), 1–25. https://doi.org/10.1186/s41239-024-00487-w

Warschauer, M., Tseng, W., Yim, S., Webster, T., Jacob, S., Du, Q., & Tate, T. (2023). The affordances and contradictions of AI-generated text for writers of English as a second or foreign language. *Journal of Second Language Writing, 62*, Article 101071. https://doi.org/10.1016/j.jslw.2023.101071