CHAPTER 13.

# AI DETECTORS CAN STOP INTERNATIONAL STUDENTS FROM PLAGIARIZING

## ✨ *EDUCATORS SHOULD PRIORITIZE CONVERSATION, TRUST-BUILDING, AND LIVED EXPERIENCE AS PART OF CRITICAL AI LITERACY*

**Priscila Schilaro Santa Rosa and Sherry Rankins-Robertson**
University of Central Florida

In early 2023, while working as a graduate assistant for the Department of Writing and Rhetoric at University of Central Florida, Priscila tested Turnitin's AI detector while developing material for a faculty event out of the WAC program. Priscila wrote an excerpt on her experiences as an international student; she asked ChatGPT to generate a similar text. To her shock, Turnitin flagged her writing as 100% AI-generated, while ChatGPT's text was marked 0%. She emailed faculty members with the subject line: "Turnitin thinks I'm a robot!" This experience began her commitment to educating folks about the effects of AI detectors for international students.

In this chapter, we address the bad idea of using AI detectors to determine academic integrity, particularly with multilingual writers, and reinforce, as stated in this collection's introduction, that "effective writing requires navigating varied rhetorical conventions, cultural differences, and increasingly multilingual and transnational communicative practices." As the authors of this text, we identify concerns that international students may face while writing in the age of AI and offer recommendations for faculty members to gain AI literacy surrounding the effects of AI detectors and provide information literacy strategies for writing students.

### CHALLENGES INTERNATIONAL STUDENTS FACE IN WRITING CLASSROOMS

Research has found international students generally experience more

DOI: https://doi.org/1010.37514/PER-B.2026.2777.2.13 

psychological distress than native-born students due to fluency issues, cultural barriers, academic and financial problems, lack of support, loneliness, and discrimination (Maringe & Carter, 2007; Sherry et al., 2010; Yeh & Inose, 2003). International students are under pressure to maintain high grades because their visas hinge on their academic performance. These bureaucratic, social, and cultural matters complicate international students' relationship to their performance in writing classrooms.

Language fluency and performance are not just a matter of pride but of academic survival. Accusations of plagiarism can have far-reaching consequences—years of hard work, familial sacrifice, and financial investment are lost. That type of pressure can turn writing assignments into potential minefields, particularly when there is a heavy emphasis on grammar and mechanics and writing is seen as a product rather than a process. From interviews with international students, Ali R. Abasi and Barbara Graves (2008) noted that strict plagiarism policies increased student anxiety and distracted them from writing. Diane Pecorari and Bojana Petrić's (2014) research showed students often don't aim to fool their teachers, but rather plagiarism occurs due to misunderstandings of what is/is not acceptable when using sources in their writing.

## AI DETECTORS IN WRITING CLASSROOMS AND CHALLENGES FOR INTERNATIONAL STUDENTS

Writing scholars know human writing can vary depending on the writer's linguistic and sociocultural background, education, and individual style (Elbow, 2013; Gee, 2005; Young, 2010). In a review of nearly two dozen articles on AI detectors' performance, Chaka Chaka (2024b) found inconsistent efficacy and concerns about reliability. Further, Debora Weber-Wulff et al. (2023) tested twelve publicly available AI detection tools and found that 20% of texts were identified as false positives (when human writing is incorrectly flagged as AI); particularly, human-authored texts translated from any language into English had lower accuracy results. Liang et al. (2023) found detectors mistakenly flag essays by non-native writers at a higher rate as AI-generated texts—revealing biases inherent in assumptions about writing. These results on publicly available tools demonstrate the issue with current detection techniques.

AI detectors calculate the probability of a machine-generated text versus a human-written one based on unevenness or variability in the lengths or styles of sentences along with vocabulary predictability (Chaka, 2024a, p. 10). This method of detecting "machine-generated" text requires making assumptions about "human-written" text, particularly the ways that texts emulate the specific style (and culture) of (standard written) texts that large-language models train

on. Weixin Liang et al. (2023) showed AI detectors can be unreliable when the (human) author is not a native English speaker—over half of ESL-written essays were misclassified as AI-generated. Ji et al. (2024) found that human reviewers struggled to differentiate between texts written by humans and those written by GenAI. Ironically, when GenAI was used to enhance the word choice in essays written by non-native speakers, those texts were less likely to be flagged as AI-generated. This contradiction challenges the premise of distinguishing "human" from "machine"-generated text.

While a recent study by Yang Jiang et al. (2024) showed no evidence of bias disadvantaging non-native writers, the authors themselves emphasize that these results cannot be generalized to classrooms or broader writing contexts. Their findings are limited using the ETS e-rater engine and perplexity measures from GPT-2 (chosen because newer models, at the time of their study, like GPT-4 were not open-source for perplexity calculations). Moreover, Jiang et al.'s study focused only on Graduate Record Examinations (GRE) essays, a very specific, timed genre produced by a high-performing, test-prepared population (2024). By contrast, classroom writing varies widely across semesters/time and student backgrounds, making it nearly impossible to replicate such purpose-built models without reintroducing bias.

Thus, there has been emerging work to improve the reliability of detection (Chakraborty et al., 2023; Seetharaman & Barnum, 2024). Moran, assistant professor of computer science at the University of Central Florida, states, "as people use AI to generate text more frequently, and models evolve, the line between AI-generated and human-generated text will continue to blur, making it harder to differentiate between the two" (2024, personal communication). This sentiment is echoed in Jahna Otterbacher (2023), which highlights how relying only on technical solutions reinforces a never-ending technological race between detectors and generators—a race that will have teachers perpetually playing catch-up and risking false positives—while distracting their pedagogical focus and further discouraging writers in the meantime. This exponential growth in the field of artificial intelligence makes texts, such as this one, perpetually behind—the moment we grapple with the changes, new technologies arrive. This technological arms race has pedagogical implications that extend beyond detection accuracy.

As Noël Ingram (2026) argues, suspicion-based responses to GenAI risk creating adversarial classrooms that disproportionately target marginalized students. Like Ingram's call for a pedagogy of trust and joy, our recommendations emphasize that the solution is not harsher enforcement or technological "AI-proofing," but building transparent and supportive environments where students—especially international and multilingual students—can learn without fear of false

accusation. Therefore, approaches that emphasize trust and context may be the safest and most reliable way to address suspected plagiarism.

## HOW FACULTY CAN HELP (INTERNATIONAL) STUDENTS BUILD CRITICAL AI LITERACY

Educating (international) students on GenAI will do more to combat plagiarism than detection—here, we place international in parentheses, as we believe all students can benefit from critical AI literacy. Faculty members should not assume that any student, particularly international students, fully understand what constitutes plagiarism or proper citation practices, specifically in first-year composition courses. We argue for clear syllabi policies on academic integrity (separate from GenAI uses) and GenAI statements. Prior to asking students to work with sources, faculty members should explain why academic integrity matters and provide clear(er) guidelines—along with discussions on AI technologies and their affordances. Intentional instruction should include incorporating AI tools in the classroom, so students have experience critically examining their affordances and limitations. For example, AI can quickly suggest synonyms (an affordance), but it may also reinforce stereotypes (a limitation).

In an effort to teach students about GenAI, we urge faculty members to develop their own AI literacies, and understand issues with AI detectors, along with having a strong pedagogical foundation of international learners' writing and be aware of existing biases in relation to these students. The MLA-CCCC Joint Task Force's "Generative AI and Policy Development" (2024) states:

> Faculty members must examine their own implicit biases and assumptions about the relationship between language and identity, taking care to avoid making negative assumptions about marginalized writers on the basis of academic writing. Literature across the disciplines has shown that international students and multilingual students who are writing in English are more likely to be accused of GAI-related academic misconduct (Tzanni; Foltynek et al.; Weber-Wulff et al.). The problem is twofold. Studies have shown that GAI detectors are more likely to flag English prose written by nonnative speakers (Liang et al.; Weber-Wulff et al.), but even faculty members who do not use AI detectors should be aware that suspicions of misuse of GAI are often due to complex factors, including culture, context, and unconscious "native-speakerism" rather than actual misconduct (Tzanni). (p. 9)

See also the MLA-CCCC *Student Guide to AI Literacy* (2024), which outlines practical strategies for faculty and students to cultivate critical awareness in their use of AI tools.

## TIPS FOR USING GENAI IN WRITING CLASSROOMS

Some practical solutions that we use to teach GenAI literacy in our classrooms include introducing GenAI as another research and writing tool in the information literacy process—with the same thoughtful considerations as any other rhetorical choice. To help familiarize writers with GenAI tools, we advocate for listing syllabus statements about GenAI and its role in the writing process, offering step-by-step support for assignments that integrate and identify GenAI uses throughout the stages of writing (consider "AI Go" and "AI No"), inviting students to analyze biased language patterns used by AI tools, and asking students to verify sources.

When students are asked to reflect on how GenAI changed or confirmed assumptions, they articulate critical thinking. Students maintain ownership of their voice while safely experimenting with technology without fear of accusations. Students are encouraged to share their thoughts on GenAI, and as a result some students have stated GenAI tools has helped them overcome struggles with idea generation and made them feel confident as writers—transforming writing from a source of stress to a process of empowerment. These steps create a supportive learning environment for international students—allowing them, as Sherry's MA mentor often taught her, to "take risks, make mistakes, and get messy" (the rules of the magic school bus) to explore writerly rhetorical moves and share concerns about academic integrity issues. Additional classroom strategies for helping international students navigate GenAI writing can be found in Alice Gruber's "AI Will Empower Non-Native English Writers to Master 'Standard Academic English'" (this collection).

## CONCLUSION

By focusing on conversation, trust-building, and methods for connecting academic work with students' lived experiences, educators can help international students move beyond compliance or avoidance. Let standards serve as a starting point for dialogue rather than an endpoint. Thus, meeting academic standards becomes less about enforcement and more about fostering a richer relationship with knowledge and the learning community. While language and math collide beautifully within GenAI, the sociocultural factors associated with communication (and its many forms) shouldn't be ignored. As educators, we should prioritize the human factors we can solve in this equation: knowledge gaps and potential stressors for student writers.

## AUTHORS' NOTE

A special thanks to Dr. Shane Wood, Writing Program Administrator at the University of Central Florida, for his thoughtful engagement with this text and topic.

## REFERENCES

Abasi, A. R., & Graves, B. (2008). Academic literacy and plagiarism: Conversations with international graduate students and disciplinary professors. *Journal of English for Academic Purposes, 7*(4), 221–233. https://doi.org/10.1016/j.jeap.2008.10.010

Chaka, C. (2024a, April 9). Accuracy pecking order – How 30 AI detectors stack up in detecting generative artificial intelligence content in university English L1 and English L2 student essays. *Journal of Applied Learning & Teaching, 7*(1). https://doi.org/10.37074/jalt.2024.7.1.33

Chaka, C. (2024b, February 6). Reviewing the performance of AI detection tools in differentiating between AI-generated and human-written texts: A literature and integrative hybrid review. *Journal of Applied Learning & Teaching, 7*(1). https://doi.org/10.37074/jalt.2024.7.1.14

Chakraborty, S., Singh Bedi, A., Zhu, S., An, B., Manocha, D., & Huang, F. (2023). *On the possibilities of AI-generated text detection*. arXiv. https://doi.org/10.48550/arXiv.2304.04736

Elbow, P. (2013). Using careless speech for careful, well-crafted writing—whatever its style. *The Journal of the Assembly for Expanded Perspectives on Learning, 19*. https://doi.org/10.7290/jaepl192flw

Fowler, G. A. (2023, August 14). What to do when you're accused of AI cheating. *The Washington Post*. https://tinyurl.com/mvrw56p6

Gee, J. P. (2005). The new literacy studies: From "socially situated" to the work of the social. In D. Barton, M. Hamilton, & R. Ivanic (Eds.), *Situated literacies: Theorising, reading and writing in context* (pp. 180–196). Routledge. https://doi.org/10.4324/9780203984963

Gruber, A. (2026). AI will empower non-native English writers to master "standard academic English": AI that reflects global Englishes usage and language diversity can support students' writing. In C. Basgier, A. Mills, M. Olejnik, M. Rodak, & S. Sharma (Eds.), *Bad ideas about AI and writing: Generative practices for teaching, learning, and communication*. The WAC Clearinghouse; University Press of Colorado. https://doi.org/10.37514/PER-B.2026.2777.2.27

Ingram, N. (2026). We should suspect our students are using AI: Instead of mistrusting our students, we should embrace a pedagogy of trust and joy. In C. Basgier, A. Mills, M. Olejnik, M. Rodak, & S. Sharma (Eds.), *Bad ideas about AI and writing: Generative practices for teaching, learning, and communication*. The WAC Clearinghouse; University Press of Colorado. https://doi.org/10.37514/PER-B.2026.2777.2.40

Ji, J., Li, R., Li, S., Guo, J., Qiu, W., Huang, Z., Jiang, X., & Lu, X. (2024). *Detecting machine-generated texts: Not just "AI vs humans" and explainability is complicated*. arXiv. https://doi.org/10.48550/arXiv.2406.18259

Jiang, Y., Hao, J., Fauss, M., & Li, C. (2024). Detecting chatgpt-generated essays in a large-scale writing assessment: Is there a bias against non-native English speakers? *Computers & Education, 217*, 105070. https://doi.org/10.1016/j.compedu.2024.105070

Klee, M. (2023, June 7). She was falsely accused of cheating with AI – And she won't be the last. *Rolling Stone*. https://tinyurl.com/yc5nvjud

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

Maringe, F., & Carter, S. (2007). International students' motivations for studying in UK HE. *International Journal of Educational Management, 21*(6), 459–475.

Mathewson, T. G. (2023, August 14). AI detection tools falsely accuse international students of cheating. *The Markup*. https://themarkup.org/machine-learning/2023/08/14/ai-detection-tools-falsely-accuse-international-students-of-cheating

MLA-CCCC Joint Task Force on Writing and AI. (2024, April). *Working paper 2: Generative AI and policy development: Guidance from the MLA-CCCC task force*. https://aiandwriting.hcommons.org/working-paper-2/

MLA-CCCC Joint Task Force on Writing and AI. (2024). *Student guide to AI literacy*. https://aiandwriting.hcommons.org/student-guide-to-ai-literacy/

Otterbacher, J. (2023). Why technical solutions for detecting AI-generated content in research and education are insufficient. *Patterns, 4*(7), 100796. https://doi.org/10.1016/j.patter.2023.100796

Pecorari, D., & Petrić, B. (2014). Plagiarism in second language writing: Is it time to close the case? *Language Teaching, 47*(3), 269–302. https://doi.org/10.1017/s0261444814000056

Seetharaman, D., & Barnum, M. (2024, August 4). There's a tool to catch students cheating with ChatGPT. OpenAI Hasn't Released It. *Wall Street Journal*. https://www.wsj.com/tech/ai/openai-tool-chatgpt-cheating-writing-135b755a

Sherry, M., Thomas, P., & Chui, W. H. (2010). International students: A vulnerable student population. *Higher Education, 60*(1), 33–46. http://www.jstor.org/stable/40784044

Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., Šigut, P., & Waddington, L. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity, 19*(1). https://doi.org/10.1007/s40979-023-00146-z

Yeh, C. J., & Inose, M. (2003). International students' reported English fluency, social support satisfaction, and social connectedness as predictors of acculturative stress. *Counseling Psychology Quarterly, 16*(1), 15–28. https://doi.org/10.1080/0951507031000114058

Young, V. A. (2010). Should writers use they own English? *Iowa Journal of Cultural Studies, 12*(1), 110–117. https://doi.org/10.17077/2168-569x.1095