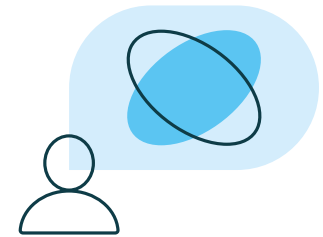


AI conversations: Handling false positives for educators

False positive: Incorrectly identifying fully human-written text as AI-generated text.

While the risk for encountering a false positive when using AI writing detection on student work is small, it is important to plan for how to approach the conversation with students. This guide shares strategies educators can consider before and after submissions to help when confronted with a false positive.



NOTE: All of our guidance rests on the strength of the relationships that instructors build irrespective of AI use or misuse. That is the most important step to help ease the difficulties these conversations might bring.

Before/during the assignment:

- **Clearly articulate what is/is not permissible, even on a particular assignment:** If possible/available, reference any institutional and/or departmental guidance. Set the expectation that if students use AI writing tools, they must provide a citation.
- **Collect a diagnostic writing sample:** A writing sample can serve as a baseline to which educators can compare. (Pro Tip: Do this as an on-demand, in-class assignment to further minimize the possibility of misconduct.
- **Employ our AI misuse tools to prepare assignments:** Use our [AI misuse rubric](#) to make sure the writing assignments are less vulnerable to misconduct. Use our [AI misuse checklist](#) to structure the writing process in ways that cut down on the threat of any unauthorized use of AI writing tools.
- **Plan ahead for the process that will be used if a question of AI misuse arises:** Set the expectation about AI misuse for students so that they aren't caught unaware. Will the educator ask the student challenge questions that might show how well they understand the content of the writing? Will the educator ask the student reflection questions designed to reveal the writing choices made in the piece? Will the educator ask the student probing process questions that might reveal gaps or issues?

After the assignment is submitted:

- **Use Turnitin's AI writing detector:** Understand [how Turnitin's AI writing detector works](#) and use it to evaluate student work. Remember that **the indicator is not a replacement for the educator's judgment**. Our goal is to provide the insights needed to make good decisions, but it is only one piece of the puzzle. Educators know their students and their work, and that is an equally important piece of data.
- **Revisit your assignment and your process:** Were guardrails in place to protect the assignment from AI misuse? Try to determine the risk of misuse on the assignment and make changes as necessary.
- **Rely on relationships with the student:** This kind of judgment should never be made without a respectful dialogue with the student. Educators will have built relationships with the student; use that as the filter for evaluating what actually happened here.
- **Compare the writing in question to the diagnostic sample:** Does the flagged writing align to the style, complexity, and sophistication of the student's previous writing? One thing to look for are distinctive phrases - idioms or expressions that seem out of place; discussing these should be part of the dialogue with the student.
- **Adopt an attitude that assumes positive intent:** After review, if the evidence isn't clear, give the student the benefit of the doubt. All the right conversations have taken place, with all the right questions asked, and there's still uncertainty, the student cannot be penalized based on that. The good news is that the experience alone should act as a powerful deterrent for any potential future misuse!

To explore more ideas about academic integrity in the age of AI, visit www.turnitin.com/resources/academic-integrity-in-the-age-of-AI

