

# The Case Against AI Detectors

With the rise of large language models (LLMs) like ChatGPT, Claude, and Gemini, companies have introduced AI detection tools that claim to help educators determine when an LLM has been used to complete an assignment. Here are four reasons the Hastings Initiative does not support the use of such tools:

---

## Unreliability

1

Performance varies across text length, genre, and hybrid texts ([Hadra et al. 2026](#)). AI detectors also often produce vastly different scores on repeated analysis of the same text, meaning educators cannot make consistent or defensible judgements based on their outputs ([Malik and Amjad 2025](#)).

---

## Bias Against Non-Native Speakers

2

A Stanford study found that AI detection tools misclassify texts written by non-native English speakers more frequently than text by native English speakers ([Liang et al. 2023](#)). As a result, they can unfairly penalize students who are non-native English speakers.

---

## Atmosphere of Distrust

3

AI detection-focused approaches can foster an environment of distrust and anxiety in the classroom, undermining educational relationships ([Giray et al. 2025](#)).

---

## No Silver Bullet

4

With educator adoption of AI detection tools, there is a growing student market for tools that “humanize” text generated by AI to bypass detection ([Masrouf et al. 2025](#)). As LLMs and humanizer tools continue to improve, detection tools cannot solve academic integrity concerns.

---

### Further Reading

- [“Managing Artificial Intelligence \(AI\) in Teaching and Learning,”](#) Bowdoin College
- [“AI Detectors Don’t Work. Here’s What to Do Instead.”](#) MIT Sloan Technology Services
- [“Pros and Cons of AI Detection,”](#) University at Albany