NO PREP

# AI-Written Text

## How to Tell If a Human Wrote It

**Includes:** Student reading, case study, differentiated questions, teacher key

Name: Period: Date:

### AI-Written Text: How to Tell If a Human Wrote It

**The New Writing Landscape**

In 2022, the release of ChatGPT introduced the world to a new reality: machines could now generate text that reads like competent, even polished human writing. Within months, AI-generated text was appearing in student essays, corporate communications, news articles, social media posts, and even published books. The volume of AI-generated content online has grown exponentially since then. Researchers estimating the proportion of web content produced by AI since 2023 face a genuinely difficult measurement problem — because AI text can closely resemble human writing, and the tools designed to detect it are imperfect. This creates a real challenge. In educational settings, teachers need to evaluate whether students are actually

**Name:** **Period:** **Date:**

# AI-Written Text: How to Tell If a Human Wrote It

### The New Writing Landscape

In 2022, the release of ChatGPT introduced the world to a new reality: machines could now generate text that reads like competent, even polished human writing. Within months, AI-generated text was appearing in student essays, corporate communications, news articles, social media posts, and even published books. The volume of AI-generated content online has grown exponentially since then. Researchers estimating the proportion of web content produced by AI since 2023 face a genuinely difficult measurement problem, because AI text can closely resemble human writing, and the tools designed to detect it are imperfect. This creates a real challenge. In educational settings, teachers need to evaluate whether students are actually learning to write or outsourcing their thinking. In journalism and research, readers need to evaluate whether content was genuinely investigated or automatically generated. In personal communications (dating apps, customer service interactions, social media), people need to understand whether they're reading a human perspective or a machine output. Developing the ability to recognize the signs of AI-generated text is a practical literacy skill for the current moment.

### What AI Writing Does Well

To spot AI writing, you first need to understand what it does well, because its strengths are as telling as its weaknesses. Large language models like GPT-4, Claude, and Gemini are trained on enormous corpora of human text. They have internalized the patterns of well-structured writing: topic sentences, supporting evidence, transitional phrases, concluding summaries. The result is text that is almost always grammatically correct, logically organized, and tonally consistent. It rarely makes the kinds of errors that human writers make: run-on sentences, sudden tonal shifts, forgetting to close a point, going off on personal tangents. AI writing is also

comprehensively neutral in a distinctive way. It tends to present multiple perspectives on any moderately controversial topic, hedge its claims with qualifications, and avoid taking strong personal positions. This isn't because AI models are philosophically balanced; it's because they were trained to avoid generating content that could be criticized as one-sided. The result is writing that is often described as 'safe,' 'corporate,' or 'textbook' in tone: technically adequate but lacking the committed voice that characterizes the best human writing.

**The Specific Tells of AI Text**

Several linguistic patterns are strongly associated with AI-generated text. First, watch for what researchers call 'hedging language' used in excess: phrases like 'it is important to note,' 'it's worth considering,' 'while there are many perspectives,' and 'ultimately' appear at rates in AI text that far exceed their frequency in typical human writing. AI models use these phrases as structural scaffolding (they signal transitions and qualifications), but a single essay that uses 'it is important to note' three times and 'ultimately' twice in the conclusion is exhibiting a statistical signature of AI generation. Second, look for the 'balanced perspective' trap. Human writers on topics they care about tend to have opinions. They may acknowledge counterarguments, but their prose usually reflects a genuine point of view. AI writing, by contrast, often presents 'on one hand... on the other hand' structures with perfect symmetry, followed by a conclusion that essentially amounts to 'both perspectives have merit.' This isn't how opinionated humans write. If an essay on a topic you know the writer cares about feels strangely balanced and noncommittal, that's a yellow flag. Third, look for transitions that are logically correct but feel mechanical: 'Furthermore,' 'Additionally,' 'Moreover,' and 'It is crucial to understand that' are all more common in AI text than in natural human prose. Human writers vary their transitions more and often use syntactic structure rather than explicit transition words to connect ideas.

**What AI Writing Gets Wrong**

Despite its surface polish, AI writing has characteristic weaknesses. One of the most reliable is what researchers call 'confident vagueness' — writing that sounds authoritative but, on close examination, doesn't actually say anything specific. You'll encounter this in sentences like 'Throughout history, humans have grappled with the complex relationship between technology and society' — a statement that is

technically true but entirely devoid of useful information. AI models are trained to generate plausible continuations of text, and sometimes 'plausible' means 'generically true' rather than 'specifically accurate.' AI writing also tends toward what you might call 'list brain' — a tendency to organize information in parallel lists, enumerate points as 'First... Second... Third,' and rely heavily on bullet-point-style structure even in flowing prose. Human writers use lists too, but they mix structural approaches. A 500-word response that consists entirely of enumerated points and subheadings, with no discursive paragraphs, is likely AI-generated. Finally, AI text rarely makes the kind of specific, personal observations that reflect genuine experience. Human writers slip in specific details, memories, unexpected examples, and personal reactions. AI writing tends to stay at a level of abstraction that is correct and applicable but impersonal — it could apply to anyone rather than to this specific person in this specific situation. The absence of specificity is one of the most reliable tells of all.

**The Limits of AI Detection Tools**

Several commercial tools promise to detect AI-generated text — Turnitin's AI detection feature, GPTZero, and others. These tools analyze statistical patterns in text to estimate the probability that it was AI-generated. In testing, they perform reasonably well at identifying clearly AI-generated text but struggle significantly in several scenarios: text that was AI-generated and then edited by a human, text written by a non-native English speaker whose writing naturally has some statistical patterns resembling AI output, and highly technical writing that, by its nature, sounds formal and structured. The false positive problem is serious. Multiple studies have documented cases where detection tools flagged clearly human-written text as AI-generated — including published work by prominent authors, college essays written by students who had never used AI, and essays by students with certain learning differences whose writing patterns triggered detection algorithms. Given these limitations, detection tools should be treated as one signal among many, not as definitive evidence.

**Reading for Voice, Reading for Specificity**

The most reliable human approach to detecting AI writing doesn't involve technology at all — it involves reading for voice and specificity. Authentic human writing, even when it's trying to sound professional or academic, has a voice. It makes choices. It

reflects genuine enthusiasm, genuine uncertainty, genuine perspective. It includes examples and details that couldn't have been generated by a model that has never actually experienced anything. When you read a piece of writing and feel like you're hearing a real person, you probably are. When you read writing and feel like you're reading a well-organized but somehow empty document, you may be reading AI. In practical terms: look for the unexpected. Genuine human writing surprises you with specific details, odd observations, and ideas that feel like they came from somewhere real. AI writing surprises you far less often — it tends toward the expected, the safe, the comprehensively reasonable. The best single question to ask about any piece of writing: does this feel like it was written by someone who actually cares? One additional tell worth noting is the characteristic ending of AI-generated text: a summary paragraph that systematically restates every point made in the piece, with a closing sentence that broadly endorses the importance of the topic. Human writers, especially strong ones, tend to end with a final specific insight, a provocative question, or an opening to further thought — not a comprehensive restatement of everything that preceded it.

**Key Information: Signs of AI-Generated Text**

| Pattern | Description |
|---|---|
| Excessive hedging language | Overuse of 'it is important to note,' 'ultimately,' 'it's worth considering' |
| Perfectly balanced perspectives | Equal 'on one hand / other hand' without genuine position |
| Mechanical transitions | 'Furthermore,' 'Additionally,' 'Moreover' in high frequency |
| Confident vagueness | Authoritative tone without specific or useful content |
| List brain structure | Heavy reliance on enumerated points and parallel lists |
| Absence of personal detail | No specific memories, observations, or unexpected examples |
| Tonal consistency | No variation in energy, no tonal shifts or personal tangents |

| Pattern | Description |
|---|---|
| False positive risk | Detection tools can incorrectly flag human writing as AI |

**Case Study: Ms. Reyes and the Essay**

**Ms. Reyes, an English teacher, received a personal essay from a student named Devon. The assignment asked students to write about a formative personal experience. Devon's essay was perfectly grammatical and well-organized, describing 'a time when I faced a challenge and grew from it.' The essay used phrases like 'this experience taught me that perseverance is important' and 'ultimately, I came to understand that resilience is a core human value.' There were no specific details about where the event occurred, no names of people involved, no dialogue, and no observations that couldn't have been written by anyone about any vague challenge. The conclusion stated that 'both the difficulty and the growth were equally valuable parts of the journey.' Ms. Reyes couldn't point to a single grammatical error or structural problem — but the essay felt hollow in a way she struggled to describe to Devon when she raised her concern.**

1. Using specific terms and concepts from the reading, identify at least three tells in Devon's essay that suggest it may be AI-generated.

2. Why was Ms. Reyes in a difficult position when trying to raise her concern with Devon, and what does this reveal about the limitations of relying solely on detection tools or explicit rule violations?

**Multiple Choice**

1. What does the term 'confident vagueness' mean in the context of AI-written text?

a) Writing that is clear and precise but lacks emotional depth

b) Writing that confidently presents one side of an argument without nuance

c) Writing that is grammatically incorrect but still sounds like it carries authority

d) Writing that sounds authoritative but contains no specific or useful information

2. Why do AI detection tools sometimes produce false positives — flagging human writing as AI-generated?

a) Detection tools only analyse sentence length, ignoring word choice, structure and word register

b) Non-native speakers and writers with certain styles can have patterns that resemble AI output

c) Detection tools are deliberately calibrated to favor AI companies

d) Human writers increasingly copy AI writing styles intentionally

3. According to the reading, what is the most reliable non-technological approach to detecting AI writing?

a) Reading for voice and specificity — looking for genuine personal perspective and unexpected details

b) Checking whether the text cites sources with verifiable URLs

c) Counting the frequency of transition words in the text

d) Running the text through several different detection tools at once and comparing all their verdicts

**Matching**

Write the letter of the correct definition on the line next to each term.

| Term | Definition |
|---|---|
| 1. ___ Hedging language | A) AI system trained on massive text datasets to generate human-like writing |
| 2. ___ False positive | B) Writing that sounds authoritative but lacks specific meaningful content |
| 3. ___ Large language model | C) Incorrect classification of real human writing as AI-generated |
| 4. ___ Confident vagueness | D) The distinctive perspective, tone, and personal quality of genuine human writing |
| 5. ___ Voice | E) Qualifiers like 'it is important to note' used excessively in AI text |

**Fill-in-the-Blank**

1. AI writing tends to present perfectly balanced perspectives followed by conclusions that 'both sides have merit,' reflecting the models' training to avoid generating _____ content.

2. The tendency of AI text to organize information in parallel numbered points and rely heavily on enumerated structure is sometimes called '_____ brain.'

3. The most reliable tell of AI writing may be the absence of _____, since AI models have never actually experienced anything and cannot generate genuinely personal observations.

**Short Answer**

Answer each question in 2-4 complete sentences.

1. The reading describes AI writing as having strengths that are 'as telling as its weaknesses.' Choose one AI writing strength described in the reading and explain how that same strength becomes a signal that the text may be AI-generated.

2. Explain the 'false positive problem' with AI detection tools and why it matters for how we should use these tools in educational and professional settings.

3. The reading ends with the question: 'does this feel like it was written by someone who actually cares?' Why might caring be a reliable marker of genuine human writing, and what are the limits of using this as a detection standard?