AI for Teachers



Saving 10 Hours a Week with Claude and ChatGPT

Sarah Whitman

AI for Teachers

Saving 10 Hours a Week with Claude and ChatGPT

by Sarah Whitman

CONTENTS

I Why I Started Using AI (and Why I Almost Didn't)

II The Three Tools I Rely On (and the Ones I Tried and Dropped)

III Lesson Planning That Finally Fits a Sunday Afternoon

IV Differentiation That Doesn't Kill You

V Feedback on Student Writing — What AI Does Well, What It Doesn't

VI Rubrics and Assessments

VII Parent Communication

VIII IEP and 504 Documentation

IX The Hard Conversation: Students Using AI

X Honest Detection (vs. Accusations You Can't Prove)

XI Teaching ABOUT AI in Your Subject

XII Your First 30-Day Teacher AI Plan

CHAPTER I

Why I Started Using AI (and Why I Almost Didn't)

It was a Sunday in November, and I was losing.

Not dramatically. Not in a way anyone would see. I was sitting at the kitchen table in my apartment in Clintonville, a neighborhood on the north side of Columbus, with a mug of tea that had gone cold two hours earlier. My stack of 112 eighth-grade argument essays — printed, because I still marked in pen back then — was sorted into three piles by class period, which was the most productive thing I had managed to do since noon. Three lesson plans for the week were open in separate browser tabs, two of them barely started. And somewhere under the essays was a folder of IEP progress notes I owed the special education coordinator by eight o'clock Monday morning. Seven of them. Each one requiring language specific enough to be useful, vague enough to be defensible, professional enough to survive a due-process review.

It was 9:47 p.m.

I had been a teacher for six years. I loved my job — I want to be clear about that, because this story doesn't end with me quitting, and everything in the next eleven chapters is written by someone who still walks into school on Monday mornings not entirely miserable. But that Sunday I was doing the math that teachers aren't supposed to do out loud, the math that you've probably done yourself in a parking lot or a bathtub or some other place where no one can see your face while you're doing it. According to a 2019 RAND Corporation survey of over 1,000 teachers across the country, the average teacher works 10.5 hours on school days and an additional 7.5 hours over the weekend. I was above average that night and not in a way that felt like an achievement.

I didn't Google "AI for teachers." That's not how it happened.

What I actually typed, at 9:52 p.m., into the ChatGPT search bar I'd opened mostly out of desperate curiosity after a colleague mentioned it in the copy room, was this: Can you help me write a present level of performance summary for a student who struggles with organizing multi-paragraph writing and has processing speed challenges?

I want to tell you about what came back. But first I have to tell you who I was before I pressed enter, because that context matters. If you're reading this as a skeptic — and I hope you are, because the true believers tend to make bad decisions about technology — you deserve to know that I was one of you.

The Eye-Roll File

I have been teaching long enough to have participated in, or been subjected to, the following professional development initiatives: interactive whiteboard training (twice, because we got new ones), a full-year

"learning styles" cohort where we redesigned every lesson to include visual, auditory, and kinesthetic components, a flipped-classroom pilot that lasted one semester before the equity issues became undeniable, and a district-wide Google Classroom rollout that consumed what felt like six weeks of faculty meetings and produced a generation of students who are genuinely very good at submitting things to the wrong folder.

I'm not being cynical. I genuinely engaged with most of these. I made kinesthetic vocabulary stations. I filmed myself explaining direct objects on a seven-minute video that maybe four students actually watched. I believed in the theory behind a lot of it. The problem was that the research behind much of it was somewhere between thin and nonexistent. The learning-styles inventory, for example — the idea that students are categorically visual, auditory, or kinesthetic learners who perform better when instruction matches their "style" — was conclusively debunked in a 2008 review by cognitive psychologists Harold Pashler and his colleagues, published in Psychological Science in the Public Interest. The research had been weak for decades. We were still doing the stations.

So when ChatGPT became a staff-room conversation topic in the fall of my sixth year, my internal response was practiced. Here we go. Another thing that was going to transform education. Another vendor presentation with stock photos of children looking delighted at screens. Another pilot program, another data collection cycle, another PD day devoted to something that would feel irrelevant by March.

"I've survived interactive whiteboards and learning-style inventories," I told my department chair, who was already experimenting with AI and kept sending me links. "I'll survive this too."

She sent me another link.

I deleted it.

9:52 p.m.

Here's the honest version of how I ended up on ChatGPT: I ran out of other options. I had been staring at a blank IEP progress note for twenty minutes. The student in question — I'll call him Marcus — was a genuinely complicated kid to write about. He had real strengths. He also had a processing speed score in the 14th percentile and a habit of writing brilliant first sentences and then abandoning the paragraph entirely, as though the sentence had used up all available fuel. I knew exactly what was true about his writing. I had no idea how to say it in the flat, official prose that IEP documents require.

So I typed the question. Not with any particular faith. More the way you try a home remedy at midnight — not because you believe in it, but because you've exhausted the things you do believe in.

The response appeared in maybe four seconds.

It was not perfect. I want to be specific about that, because I think the stories where AI immediately produces something miraculous are the ones that make other teachers feel like they're missing something when reality is messier. The draft it gave me had two sentences I would never use, a phrase that was too clinical even for an IEP, and a Lexile-level reference I'd have to verify independently. But embedded in the middle of that

response was a formulation of Marcus's present performance that was genuinely accurate, professionally worded, and — this was the part that stopped me — expressed something I had been circling for weeks without landing on.

"[Student] demonstrates emerging organizational skills in extended writing tasks and benefits from explicit pre-writing scaffolds; current performance reflects a gap between verbal ideation and written production that is consistent with documented processing-speed differences."

I had not written that sentence. But it was true. And it took me forty-five seconds to edit it into something I could sign my name to.

I sat with that for a moment.

Then I opened the next IEP note and typed another question.

What Changed, and What Didn't

I finished the seven progress notes in under an hour. Not because AI wrote them — I wrote them, I reviewed every word, I changed things that were wrong or generic — but because I had somewhere to start. That turns out to be an enormous amount of the battle in documentation work. The blank page is its own kind of tax on your time and energy, and when you're doing it at 10 p.m. after a full week of teaching, that tax compounds fast.

The lesson plans and the essays didn't get finished that night. I'm not going to pretend I discovered a productivity miracle in a single Sunday session. What happened was smaller and, in the long run, more durable. I became curious instead of dismissive. That's all. Curiosity is enough.

Over the next four weeks I used AI tools about a dozen times, each time for something narrow and specific — a quiz on The Outsiders, a parent email I couldn't figure out how to open, a rubric for a podcast project I'd invented on a Tuesday with too much confidence and not enough planning. Each time the output needed editing. Most of the time the editing was fast. Once the output was genuinely bad and I scrapped it. But I kept going back, the way you keep going back to a hardware store that has an annoying layout but usually has the thing you need.

The moment AI stopped feeling like a gimmick — the real moment, not the Netflix-special version — was a Thursday in January when I was planning a unit on argument writing and I realized I had not spent Sunday night on lesson planning at all. I had done it in pieces across three lunch breaks, working from a prompt template I'd built over the previous month, and the plans were better than my Sunday-night plans had usually been. More specific. Better questions. A sequence that actually built on itself instead of listing activities in roughly chronological order and hoping for the best.

I sat at my desk during my actual prep period and read a novel for twenty minutes.

That had not happened in six years.

What I'm not going to tell you is that AI fixed everything, or that the workload is now manageable in some global sense, or that you should feel bad for not having adopted this sooner. Teacher working conditions in this country are genuinely bad — RAND's surveys have tracked declining job satisfaction and increasing attrition for years, and the problem is structural, not technological. No app solves large class sizes or underfunding or the quiet expectation that professionals will absorb unreasonable demands without complaint.

What I can tell you is that I found, on average, between eight and ten hours per week that I'd been spending on tasks that AI can do most of, or draft most of, or start for me. That's real time. It doesn't come from cutting corners on the things that matter. It comes from stop doing by hand the things that don't require hands.

The chapters ahead show you exactly where those hours come from and how to get them. But before we go anywhere, we need to know where your time is actually going, because I'd be willing to bet it's not where you think it is.

The Time Audit: One Week, No Guessing

Most teachers, when asked where they spend their time, describe their time the way they wish it worked. Planning: one hour. Grading: one hour. Communication with families: minimal. This is almost never accurate. The RAND 2019 data found that teachers systematically underestimate the hours they spend on non-instructional work — documentation, email, prep, coordination — and overestimate time spent in direct student contact.

You don't need a research study. You need one week of honest tracking.

Here's how to do it. Keep a notes app or a small paper notebook with you for five school days — Monday through Friday — and log every task you do outside of instructional time (before school, during prep, at lunch if you're working, after school, evenings, weekends). Don't write what you planned to do. Write what you actually did, and for how long. Be boring and specific. "Emailed parent re: missing work, 12 min" is useful. "Admin stuff" is not.

At the end of the week, sort your entries into these six buckets:

Lesson planning and curriculum design. Any time you spent creating, modifying, or finding materials for upcoming instruction.

Grading and feedback. Reading student work, writing comments, entering scores.

Documentation. IEPs, 504 notes, progress reports, any formal record-keeping.

Communication. Emails, phone calls, text-app messages to parents, coordination with colleagues or specialists.

Meeting and professional development time. Any scheduled time that was not classroom instruction.

Everything else. The category you'll underestimate until you see it. Printing, making copies, tracking down supplies, troubleshooting tech, hunting for that one PDF you know you saved somewhere.

Add each bucket up. The totals are often surprising in ways that are both demoralizing and clarifying. When I ran this audit for the first time — not that November Sunday, but a few months later, after a colleague walked me through the exercise — I found that I was spending 4.5 hours per week on parent communication alone. Not 30 minutes. Four and a half hours. I was spending 6 hours on lesson planning that could have been 90 minutes with the right system. I was spending almost no time on grading during school hours and compressing the entire load into Sunday evenings, which is why Sunday evenings felt like a punishment.

Once you can see the shape of your week, you know where to start. Chapters three through eight are organized around the buckets that eat the most time for most teachers — planning, differentiation, feedback, documentation, communication. You don't have to implement everything. You just have to find your biggest bucket and start there.

Run the audit this week. Seriously, this week, before you read the next chapter. The specific numbers you find will make everything else in this book land differently, because you'll stop reading about a teacher named Sarah in Columbus and start reading about yourself.

Your hours are in there somewhere. Let's go find them.

CHAPTER II

The Three Tools I Rely On (and the Ones I Tried and Dropped)

There is a version of this chapter where I list every AI tool available as of right now, rank them on seventeen metrics, and tell you which one scored highest. I am not writing that chapter. By the time you read this, three of those tools will have released major updates, two will have merged or been acquired, and at least one will have quietly shut down. I know because I spent the better part of a year chasing that list, and it cost me more hours than it saved.

What I want to give you instead is what actually worked: two tools, used consistently, set up once in a way that makes them immediately useful for a working teacher. That's it. Two tabs open on a school-issued Chromebook in Room 214 of Mifflin High School, Columbus, Ohio — and a combined weekly planning time that dropped from somewhere around thirteen hours to something that now fits between Saturday morning coffee and my daughter's noon soccer game.

But first, the ones I dropped.

The Graveyard of Good Intentions

When I finally gave in and started experimenting with AI in January of my seventh year, I did what any overwhelmed person does: I tried everything. I had a Google account, so I started with Gemini. It was fine. Genuinely fine, in the way that a slightly-too-warm classroom is fine — tolerable, never quite comfortable. When I asked it to help me write a progress-update letter for a parent conference, it produced something polished but oddly distant, like a letter written by someone who had read a lot of teacher letters but never actually felt nervous before a parent meeting. I used it twice, then stopped opening it.

I tried Microsoft Copilot because it was embedded in the tools my district already paid for. The integration with Word and Teams was genuinely smart. But Copilot kept being slightly wrong — not embarrassingly wrong, just wrong enough that I had to verify everything twice, which defeated most of the time savings. (Your mileage may vary. If your district is deep in the Microsoft ecosystem, Copilot's document integration might be worth the tradeoff.)

I tried Perplexity for research-style tasks. I tried a specialized "teacher AI" tool that a vendor pitched at our August professional development day. I spent forty minutes setting up an account for a "curriculum design assistant" that three teachers in my building were supposedly raving about — and then I never went back because it kept producing lesson plans with standards citations that didn't exist.

Here is what I learned from all of that: the tool you use inconsistently gives you nothing. Fluency matters. The more you use a particular AI, the better your prompts get, the better you understand its quirks, and the faster the whole workflow runs. Spreading your time across six platforms keeps you permanently in beginner mode on every single one.

I settled on two. I still use two.

Claude: The One I Trust with the Hard Stuff

Claude is built by Anthropic, a company founded by researchers who spent years thinking seriously about AI safety before they built a product. Whether you find that reassuring or irrelevant, you'll notice it in how the tool behaves: Claude is careful, thorough, and genuinely good at holding a lot of context at once.

That last part is why I use it for the tasks that require depth.

When I paste in three pages of a student's argument essay and ask for detailed structural feedback, Claude tracks the whole essay simultaneously — it notices that the rebuttal paragraph on page two contradicts the claim in paragraph five, and it tells me. When I'm building a five-day unit plan and I want the exit tickets to actually connect back to the learning objectives I listed at the top, Claude does that. Consistently. Without me having to re-paste the objectives every time.

I also trust Claude when I need to write something that carries emotional weight. Parent communications that aren't just informational — the ones where you need to say something hard without it landing wrong — those go to Claude. I give it my bullet points, the relationship context ("we've had one previous concern conversation in October, it went okay"), and the tone I'm going for. What comes back requires real editing, but it reads like a letter I might actually write on a good day, not a template with the student's name pasted in.

The other thing Claude does unusually well is follow complex, multi-part instructions. If I say "write this at a 6th-grade reading level, use Ohio's CC.8.W.1 standard as the stated objective, include at least two places where students have to make a personal connection, and keep it under two pages," it does all four of those things. Not always perfectly, but reliably. That reliability is underrated. It means I can write a detailed prompt once, get a useful draft, and spend fifteen minutes on genuine editing rather than thirty minutes fixing structural problems.

For the 2025–2026 school year, Claude Pro runs $20 a month. There is a free tier, which I used for about six weeks before deciding the speed limits and message caps were slowing me down more than $20 could justify. Make that call yourself based on how often you plan to use it.

ChatGPT: The One I Use When I'm in a Hurry

I know there are teachers who use ChatGPT for everything and teachers who swear by Claude for everything, and honestly both camps have reasonable arguments. But in my workflow, they do different jobs.

ChatGPT is fast. Not just fast to respond — fast to interact with. When I need a quick brain dump of possible discussion questions for The Outsiders, when I want to reformat a vocabulary list into a two-column table, when I'm staring at a quiz I drafted and I want ten possible multiple-choice distractors for question four in about ninety seconds — that's ChatGPT. The conversational rhythm is snappier. It's easier to fire off a short follow-up and get an immediate tweak.

ChatGPT is also better, in my experience, for creative brainstorming. Ask it for fifteen possible hook activities for a lesson on the persuasive techniques in advertising, and you'll get fifteen actual options, many of them genuinely playful. Some will be terrible. Some will be things I've done before. But there will usually be two or three that make me think oh, that's good — and that's exactly what you want from a brainstorming partner.

ChatGPT-4o, which became the default model in 2024, handles most classroom tasks easily. It also handles images — so if you have a printed graphic organizer you want to reproduce digitally, or a handwritten rubric from a colleague you want typed up, you can photograph it and paste it in. That function alone has saved me a painful amount of retyping.

ChatGPT Plus is also $20 a month. There is a free tier, which runs on a slightly less capable model but is entirely usable for low-stakes tasks. I kept free ChatGPT for probably four months before upgrading.

The Comparison You Actually Need

Here's the summary I wish someone had handed me on day one.

| | Claude (Pro) | ChatGPT (Plus) | |---|---|---| | Cost | $20/month; free tier available | $20/month; free tier available | | Best for | Long-form drafting, complex multi-part instructions, nuanced feedback, IEP documentation, parent letters | Rapid brainstorming, quick reformatting, visual input (images), short iterative tasks | | Context window | Very large — holds an entire unit plan or long document in memory during a conversation | Large — sufficient for most classroom tasks; earlier versions struggled more with long docs | | Tone and voice | More careful, more measured; better at matching a specific register when you describe it | More casual and fast; good at mimicking stated styles but less likely to ask clarifying questions | | Classroom limitations | Slower conversational back-and-forth; can over-explain; occasional resistance to very short outputs | More prone to confident errors on specific factual content; brainstorm outputs require more filtering | | Student data | Do not paste identifiable student information into either the free or paid consumer versions without checking your district's data policy (see Chapter 8) | Same | | Mobile app | Yes | Yes | | Best first task to try | A parent progress-update email or a present-level-of-performance summary | A set of discussion questions or a vocabulary activity for a lesson you're teaching next week |

Neither tool is always better. Neither is always right. The goal is to know quickly which one to open, which is what the decision framework at the end of this chapter is for.

The Setup Step That Changed Everything

About three weeks into using both tools, I noticed I was typing the same background information at the start of nearly every conversation. I teach 8th-grade ELA at a public school in Columbus, Ohio. My students range from 5th- to 11th-grade reading levels. I follow the Common Core State Standards. Over and over. It was eating three minutes off every session, and the outputs were still generic because the AI didn't know anything real about my context.

Then I learned about system prompts — and everything got noticeably better.

A system prompt is a block of context you give the AI once, before any actual task, that shapes everything it produces in that conversation. Think of it as the briefing you'd give a very capable new colleague on their first day. Instead of re-explaining who you are and what you need every time, you write it once, save it somewhere, and paste it at the top of any new conversation.

Here's the one I use. It's not long:

You are working with an 8th-grade ELA teacher at a public middle school in Columbus, Ohio. The school serves approximately 850 students; roughly 40% qualify for free or reduced lunch. My classes average 28 students. I follow Common Core State Standards, specifically the grade 8 ELA standards. My students' reading levels range roughly from 5th to 11th grade. I frequently have ELL students and students with IEPs or 504 plans. When helping with instructional materials, please default to a 7th–8th grade reading level unless I specify otherwise. When writing communications, use a warm, professional tone — not corporate. When drafting documents I will sign, write in first person. Ask me clarifying questions before starting if the task is ambiguous.

That's it. Ninety-four words. Paste that at the top of a Claude or ChatGPT conversation before you type your actual request, and the outputs shift immediately. The lessons sound like they're built for your students, not a hypothetical class. The parent letters sound like they're from you, not a district template. The differentiation suggestions actually account for your range.

Save your version of this in a note on your phone, a sticky in your browser bookmarks, a Google Doc titled "AI context" — wherever you'll actually find it in ninety seconds when you open a new browser tab on a Sunday afternoon. Customize it with your actual grade, your actual school context, your actual standards. If you teach AP Chemistry at a suburban high school in Naperville, the AI should know that. If you teach 2nd grade in a dual-language program in San Antonio, the AI should know that too.

The research on this tracks. A 2024 study from Stanford's Graduate School of Education examining teacher AI adoption found that teachers who established consistent contextual prompts early in their use reported significantly higher satisfaction with AI-generated output — and, critically, spent less time editing it — than those who started each session from scratch. The setup cost is one-time. The benefit compounds every single session.

Which Tool to Open: A Simple Framework

You are standing in front of your laptop on a Sunday. You have forty-seven minutes before you have to start dinner. You need to get something done. Which tab do you open?

Open Claude when:

- The task requires more than two paragraphs of output
- You need the output to hold together as a coherent whole (unit plan, multi-section rubric, IEP narrative)
- The communication has emotional stakes (anything that involves a parent concern or a student in crisis)
- You're going to paste in a long document for it to analyze or respond to
- You want it to produce something that sounds like you wrote it on a careful day

Open ChatGPT when:

- You need ideas fast and you want volume (ten hook options, twenty discussion questions, a list of exit ticket stems)
- You're reformatting something that already exists — changing a paragraph into bullet points, building a table, making a vocab list into a matching activity
- You want to photograph something physical and have it turned into text
- You're iterating quickly — changing the Bloom's level of a question, adjusting the length, trying a different format — and you want the back-and-forth to feel responsive
- You need it done in under three minutes and done is more important than perfect

There will be plenty of tasks where either tool works. That's fine. Pick one, run with it, don't agonize. The difference in output quality for a routine quiz question or a quick bellringer activity is genuinely small. The habit of opening a tool and using it matters more than the perfect tool selection in every case.

What I want you to avoid is the loop I was stuck in for months: opening three different tools, spending eight minutes comparing their outputs, deciding none of them are quite right, and closing the laptop without producing anything useful. That is not a tool problem. That is a decision problem. The framework above is designed to make the decision in about ten seconds so you can spend the rest of your time actually working.

A Word on Chasing the New Thing

Every few weeks, something launches. A new model, a new feature, a new competitor. My colleague Darnell, who teaches 10th-grade history down the hall, has tried approximately fourteen AI tools in the past eighteen

months. He is a genuinely early-adopter type, and he finds the testing interesting in its own right, and that is completely valid. But he has not systematized any of them, and his planning time looks basically like it did before he started.

The teachers I've talked to who actually changed their numbers — who got real hours back — are mostly people who picked two tools, got good at them, and stopped. Not because the other tools are bad, but because depth beats breadth. The prompt that works reliably is more valuable than the flashiest new interface.

So here's the thing I'll ask you to do before you move on to the next chapter: pick the one task in your week that eats the most time and causes the most low-grade dread. Not grading — we'll get to that. Something in the planning and preparation category. A recurring parent newsletter. The quiz you write from scratch every Friday. The lesson plan for the unit you have coming up in three weeks.

Open Claude or ChatGPT. Paste in your system prompt. Type a description of that task. See what comes back.

Don't grade it yet. Don't compare it to the other tool. Just read it, notice what's actually usable, and think about how long it would have taken you to produce the same rough draft by yourself. That number — the difference between what just happened and what used to happen — is the entire argument for everything that follows.

CHAPTER III

Lesson Planning That Finally Fits a Sunday Afternoon

It is 6:47 on a Sunday evening in October. The lesson plan template on your laptop has been open since 2 p.m. You have typed and deleted the same objective three times. Your phone has a text from your mother you haven't answered, there's a load of laundry spinning, and you still have five days of instruction to design before 7:45 tomorrow morning.

I know this evening. I've lived it more times than I want to count.

What changed for me wasn't motivation or organizational systems or a new app with a pastel color scheme. What changed was that I stopped treating lesson planning as a solo creative act and started treating it as a collaboration — one where I show up with the expertise, the kids, and the context, and the AI shows up with a first draft so I'm not staring at a blank document at 6:47 on a Sunday.

This chapter is that workflow, start to finish. I'll show you the exact prompts I use, a representative slice of what comes back, and the editing pass that turns AI output into something that actually sounds like me. By the end, planning a five-day unit should take you under an hour — probably closer to 45 minutes once you've run the system a few times.

Start With the Standard, Not the Theme

Here's a mistake I made constantly in my first years: I planned around themes and activities and then reverse-engineered a standard to match. "We'll read The Outsiders and then find a standard that fits." It produced lessons I loved and students who sometimes met the objective by accident.

AI planning works much better when you do the opposite. Start with the standard — the actual CCSS or state-level language — and build forward from there. Not because the standard is holy, but because Claude and ChatGPT both perform noticeably better when you give them a concrete target to write toward.

For this example, I was building a persuasive writing unit for my 8th-grade ELA class in late October. The anchor standard:

CCSS.ELA-LITERACY.W.8.1 — Write arguments to support claims with clear reasons and relevant evidence.

That's it. Four weeks of teaching squeezed into one sentence. Here's what I did with it.

The First Prompt: Context Is Everything

Before I ask for anything, I load my system prompt — the reusable description of my teaching context I set up in Chapter 2. It lives as a saved note I paste at the top of every new chat. It looks like this:

> *You are helping an 8th-grade ELA teacher at a mid-sized public school in Columbus, Ohio. Class sizes average 28 students. The school follows CCSS standards. Students represent a wide range of reading levels (roughly 5th–10th grade equivalency). About 30% have IEPs or 504 plans. The teacher's style is conversational and discussion-heavy; she uses Socratic seminar frequently. Units run Monday through Friday, five class periods of 50 minutes each.*

That single block of context eliminates about four rounds of clarifying back-and-forth. Claude stops giving me generic suggestions designed for a hypothetical class and starts giving me things that have a fighting chance in room 214.

With that context loaded, here's the prompt I sent:

My prompt:

> *I need a five-day unit plan for 8th-grade ELA built around CCSS.ELA-LITERACY.W.8.1 (argumentative writing). The central writing task is a 4–5 paragraph argumentative essay. I want the unit to feel relevant to students, so let's anchor it to a debate students actually care about — something school-related or current enough that they'll have genuine opinions. Give me: a unit overview with a driving question, daily objectives for Days 1–5, a materials list, two or three discussion questions per day, and an exit ticket for each day. Format it clearly. Don't give me a full lesson plan — just the planning skeleton.*

Notice what I did not say: I did not say "make it engaging" or "make it creative" or "make it rigorous." Those words are meaningless to an AI. What I gave it was a format specification (five days, specific components), a scope constraint (planning skeleton, not full plan), and a student-relevance anchor (something they actually care about). Concrete inputs produce usable outputs. Vague inputs produce word salad.

What Came Back

Here is a condensed version of the actual output I received. I've kept the structure intact but trimmed the discussion questions to one per day so this chapter doesn't become a wall of text.

Unit: "Your Move" — Making Arguments That Actually Change Things Driving Question: When is an argument strong enough to make someone change their mind?

Materials: Mentor texts (2–3 short op-eds at varied Lexile levels), student argument graphic organizer, claim/evidence/reasoning (CER) reference card, peer review checklist, essay rubric

Day 1 — What Makes an Argument, Anyway? Objective: Students will distinguish between an opinion and an evidence-based argument by analyzing two brief texts.

Discussion question: "Think of the last time someone actually changed your opinion about something. What did they say or show you that worked?"

Exit ticket: Write one sentence — your claim — on a topic from a provided list. No evidence yet. Just the claim.

Day 2 — Finding Evidence That Actually Does Something Objective: Students will identify and evaluate at least two pieces of evidence relevant to their chosen claim, using the CER framework.

Discussion question: "What's the difference between evidence that proves something and evidence that just agrees with you?"

Exit ticket: For your claim from Day 1, find one piece of evidence from the provided source packet. Write it down and explain in one sentence why it supports your claim (not just what it says).

Day 3 — Building the Structure Objective: Students will draft a complete outline for their 4–5 paragraph argument essay, including a counterclaim paragraph.

Discussion question: "Why would you want to bring up the other side's best argument in your own essay? Isn't that giving away points?"

Exit ticket: Write your counterclaim sentence and your rebuttal sentence. Two sentences total.

Day 4 — Draft Day Objective: Students will produce a complete rough draft with introduction, at least two body paragraphs, a counterclaim/rebuttal paragraph, and a conclusion.

Discussion question: "What's the hardest part of your draft right now? What do you actually need help with?"

Exit ticket: Circle one sentence in your draft that you feel least confident about. Write one question you have about it.

Day 5 — Peer Review and Revision Objective: Students will apply the CER framework to evaluate a peer's draft and identify at least one structural revision to make in their own essay.

Discussion question: "What's the difference between feedback that helps and feedback that just tells someone what's wrong?"

Exit ticket: List one change you made to your draft based on peer feedback and explain why you made it.

That took four minutes to generate. Four minutes.

Is it perfect? No. I'll get to that. But here's what it gave me that I didn't have to do myself: a coherent arc across five days, discussion questions that build on each other (notice how Day 5's question about feedback echoes Day 1's question about persuasion), and exit tickets calibrated to what students actually produced that day rather than a generic "write three things you learned."

The topic it chose — open-ended, student-generated argument topics with a provided list — was fine but not quite right for October in my building. I had something more specific in mind. That's where the editing pass comes in.

The "Make It Mine" Pass — Three Steps, Eight Minutes

The fear that AI lesson plans feel generic is real, and it's also fixable. Generic output comes from generic prompts and from treating AI output as finished product. Neither is inevitable.

Here's the three-step editing pass I run every single time.

Step 1: Swap the topic anchor (two minutes).

The unit topic the AI chose was deliberately open — a list of student-selected arguments. That works for some classes. For mine in October, I already knew I wanted to use the Columbus City Schools cell phone policy debate, which had been in the news locally and which my students had feelings about. I changed the driving question from the AI's version to:

"Should Columbus City Schools ban cell phones during the school day — and what would it actually take to convince the school board?"

That one swap made every single discussion question sharper because now students were arguing about something real, not a hypothetical. I also swapped the sample op-eds in the materials list for two actual pieces: a 2023 Columbus Dispatch editorial and a piece from a student newspaper at a district high school.

Two minutes. Done.

Step 2: Read the objectives out loud (three minutes).

This sounds too simple to be useful. It isn't. I read each day's objective out loud and ask myself: could I actually measure this? "Students will identify and evaluate at least two pieces of evidence" — yes, I can look at a graphic organizer and count two pieces and check whether they have a reasoning sentence. "Students will understand the importance of claims" — I cannot measure that, and any objective with "understand" or "appreciate" as the verb should be rewritten on the spot.

Day 2's objective survived. Day 3's survived. Day 4 needed a tweak: the AI wrote "students will produce a complete rough draft," which is measurable but doesn't specify quality. I changed it to "students will produce a rough draft of at least 250 words that includes all four required sections." Specific. Checkable. Tells me at 4:30 p.m. whether students hit it or not.

Three minutes to read, notice, and fix.

Step 3: Add one personal touchpoint per day (three minutes).

This is the step that makes a plan feel like mine rather than anyone's. Each day, I add one small specific thing from my own classroom history: a book I know students love that contains an argument worth examining, a student question from a previous year that opens up exactly the right confusion, a routine my class already has (we start every discussion with what I call a "Prove It" challenge — I make a claim, they have thirty seconds to try to agree or disagree using only what they can see in the room).

Day 1, I added: Open with the "Prove It" challenge. Claim: 'This classroom is ready to argue.' Students: thirty seconds.

I know that will land. I know it because I've done it eighteen times in six years and it never fails to get even the quietest kids talking. The AI doesn't know that. I do. That's the collaboration.

Eight minutes. The plan now sounds like something that lives in my building, in October, in room 214.

A Note on What to Ask For Next

One thing I underused for my first six months: follow-up prompts within the same conversation. Once Claude has your context and your unit skeleton, you can keep building in the same window. After I had the five-day skeleton, I sent:

> *Based on this unit plan, write four multiple-choice discussion-question options for each of the Day 1 and Day 2 exit tickets, appropriate for 8th graders. Include one plausible-but-wrong distractor per question.*

Three more minutes. Exit ticket options I'd have spent 25 minutes writing myself.

Then:

> *Now write three versions of the Day 1 discussion question at different Bloom's Taxonomy levels — one at "remember/understand," one at "analyze," one at "evaluate." Keep them on the cell phone policy topic.*

Two more minutes. I have differentiated entry points for discussion without building them from scratch.

The conversation becomes a workspace. You don't have to start over every time you need something new.

Your Plug-and-Play Prompt Template

Below is the template I actually use. Copy it, fill in the brackets, and paste it into Claude or ChatGPT with your system context at the top. You'll have a unit skeleton in under five minutes.

UNIT PLANNING PROMPT — TEMPLATE

> *[Paste your system context block here.] I need a five-day unit plan for [grade level] [subject]. The anchor standard is [paste exact standard language]. The central student task at the end of the unit is [describe the product — essay, project, presentation, etc.]. Anchor the unit to [topic/text/current event/student-relevant hook]. My students will have genuine opinions about this. Please provide: — A unit overview with a driving question — Daily learning objectives for Days 1–5 (written with measurable action verbs) — A materials list (keep it simple — assume no budget) — Two or three discussion questions per day that build on each other — One exit ticket per day tied directly to that day's objective Format it cleanly. Do not write full lesson scripts — I want a planning skeleton I can build on.*

That's it. No magic words. No secret formula. The bracket-filling is the thinking; the prompt is just how you hand it to the tool.

The Honest Benchmark

Let me be precise about time, because "it saves so much time" is meaningless without numbers.

Before I had this system, a five-day unit plan took me somewhere between three and four hours on a Sunday. That included writing objectives, sourcing or creating discussion questions, building exit tickets, and hunting for mentor texts or materials. On bad weeks — the ones where I also had grading — I'd skip the exit tickets or write discussion questions so thin they weren't really questions at all.

Now, with the system running, here's my actual Sunday breakdown:

- Load context and send the main prompt: 5 minutes
- Read output, run the three-step editing pass: 8 minutes
- Send follow-up prompts for any components I still need (additional questions, exit ticket variations, a materials tweak): 10–15 minutes
- Final read-through, copy into my school's lesson plan template: 10 minutes

Total: 33–38 minutes for a five-day unit skeleton that's ready to build from.

If I also need to find or adapt a mentor text, add 15 minutes. If the unit involves media — a documentary clip, a podcast segment — add another 10. On a fully loaded Sunday where I'm planning two different preps, I'm still done in under 90 minutes total. That used to be a Sunday afternoon.

A 2023 RAND Corporation study on teacher working conditions found that elementary and secondary teachers in the U.S. reported working an average of just under 10 hours per day during the school year, with planning and preparation consistently ranking among the top time drains outside instructional hours. I'm not citing that to make you feel vindicated. I'm citing it because the math matters: if planning drops from four hours to 45 minutes, that's three hours and fifteen minutes back in your week, every week, for the rest of the year. That's not a rounding error. That's a different life.

What This Doesn't Fix

Two things to say plainly.

First, AI planning saves preparation time, not thinking time. You still have to know what 8th-grade argumentative writing should actually look like. You still have to notice when the AI gives you an objective that sounds measurable but isn't, or a discussion question that would work great for 11th graders and fall completely flat for 13-year-olds. The tool amplifies your judgment; it doesn't replace it. A new teacher with shaky content knowledge will get mediocre plans from this system. An experienced teacher will get strong ones.

Second, the first plan you generate won't be your best. Mine wasn't. My first few AI-assisted units had a slightly uniform feel — the pacing was competent but not quite right for my class's particular rhythm, and the discussion questions were a little too clean, a little too answerable. I adjusted. I saved prompts that worked and deleted the ones that didn't. After about six weeks, the system felt like mine rather than like something I was borrowing.

That's the real investment: not Sunday afternoons anymore, but a few weeks of noticing what the tool gets right and what it doesn't. That's a fair trade for three hours back every single week.

Six-forty-seven on a Sunday night. You have plans due Monday. Open the laptop, paste your context, send the prompt.

You'll be done by 7:30.

CHAPTER IV

Differentiation That Doesn't Kill You

Let me be honest with you about something.

For three years, I had a folder on my desktop labeled "DIFF MATERIALS." Inside it lived half-finished tiered worksheets, a partially leveled article about the Civil Rights Movement, and a set of sentence frames I started making for my ELL students during a planning period in October and never finished. Every time a district walkthrough was coming, I'd open that folder, feel guilty, close it, and print the same one-size worksheet I'd been using all along.

Differentiation is the thing most teachers I know privately fail at — not because they don't care, but because doing it right, from scratch, for every unit, is genuinely impossible. Not hard. Impossible. When a document from your district's instructional framework tells you to provide three to five reading-level tiers for "all major texts," the person who wrote that has never sat down to actually produce five versions of a 600-word nonfiction article on a Thursday night after a parent meeting ran long.

Here's what changed for me. Not my work ethic. Not a better planning system. AI.

The Actual Problem with Differentiation

Teachers don't skip differentiation because they're lazy. They skip it because of math.

A single well-done tiered reading — the kind where you've genuinely adjusted sentence structure, vocabulary load, and background scaffolding rather than just changing a few words — takes most teachers 45 minutes to an hour per tier. If your district expects three tiers, that's two to three hours per text. Multiply that by the eight to twelve major readings you might assign in a semester, and you're looking at 24 to 36 hours of labor that falls completely outside lesson planning, feedback, and everything else. It doesn't fit. Nothing gets cut; differentiation does.

I taught 27 students in my fourth-year class. Six were identified as reading significantly below grade level. Four were English Language Learners at different proficiency stages. Three had IEPs with reading accommodations. And two were in our district's high-ability program that technically required extension materials. That's fifteen students — more than half my class — whose needs I was supposed to address with materials I didn't have time to make. I gave everyone the same packet and told myself I'd scaffold during discussion. Sometimes I did. Often I didn't.

This is the reality most differentiation mandates ignore.

AI doesn't fix the mandate. It fixes the time.

One Document, Four Outputs, Under Ten Minutes

Here's the workflow I now use every time I introduce a new reading in my class.

I start with a single source text — whatever I'm actually teaching. Let's say it's a nonfiction article. I paste the full text into Claude. Then I run four separate prompts in the same conversation, which means Claude holds the source in its context window and I don't have to re-paste anything.

The four prompts produce:

- A simplified version adjusted for lower-Lexile readers
- A scaffolded task card with vocabulary support for the original text
- Sentence frames for ELL students completing the written response
- An extension challenge for advanced readers

Total elapsed time, once you've done it twice: eight to twelve minutes. The first time you try it, budget twenty, because you'll want to read everything carefully and make a few edits. That's still a fraction of what tiering used to cost me.

Let me show you exactly how this looks with a real example.

The Microplastics Article: A Real Walkthrough

Last October, my 8th graders were reading a science-themed nonfiction article on microplastics — tiny plastic particles under five millimeters in size that accumulate in oceans and freshwater systems. The original article ran about 620 words and sat at roughly a 960 Lexile level, which is appropriate for on-grade-level 8th graders but genuinely difficult for students reading two to three years below grade level.

This is what I typed into Claude:

"Here is a 620-word nonfiction article about microplastics. [pasted article] Please rewrite this article for students reading at approximately a 5th-grade level (around 600–700 Lexile). Keep the key facts and main argument intact. Simplify sentence structure, replace tier-three vocabulary with more accessible alternatives, and add brief explanations where background knowledge is assumed. Do not use bullet points — keep it as flowing prose so it reads like an article, not a worksheet."

Two things to notice about that prompt. First, I gave Claude a specific Lexile target range, not just "make it easier." Vague instructions produce vague results. Second, I specified the format — prose, not bullets — because I wanted the tiered version to feel like a real article, not a dumbed-down worksheet. Students notice the difference, and it matters for their dignity.

Claude returned a version that genuinely worked. It replaced "bioaccumulation" with "building up in the bodies of animals over time." It broke a long subordinate clause sentence into two shorter ones. It added a parenthetical explanation the first time it mentioned the Great Pacific Garbage Patch. The facts stayed accurate.

Then I typed a follow-up prompt in the same conversation:

"Now create a scaffolded task card for students who will read the original article. Include: three pre-reading vocabulary words with student-friendly definitions, two 'stop and think' questions embedded at natural pause points in the text, and a sentence starter for the written response: 'The author argues that microplastics are a serious problem because ___.' Keep the language warm and encouraging, not clinical."

Sixty seconds later I had a task card I could print. The vocabulary it chose — accumulate, filter, contaminate — were exactly the right tier-two words to front-load. The stop-and-think questions landed at the right moments. I moved the second question slightly earlier in the text and changed one word in the sentence starter. That was my entire editing pass.

For the ELL students, I added a third prompt:

"My class includes English Language Learners at intermediate proficiency. Please create a set of eight sentence frames they can use when writing a paragraph response to this article. Frames should scaffold a complete argument — not just a topic sentence, but evidence integration and a closing statement. Include frames at two levels: one more supported set (fill-in-the-blank style) and one less supported set (sentence openers only)."

The output gave me exactly what I asked for. A frame like "One piece of evidence the author uses is , which shows that " is more useful to an intermediate ELL student than a blank page — and it's also more useful than a worksheet that spoon-feeds so much structure that no real thinking happens. The two-level approach meant I could hand different students different versions without making it obvious who got which.

Finally, the extension prompt:

"For students who need a challenge beyond the original article, create a one-page extension task. It should ask students to do one of the following: (1) evaluate the author's use of evidence by identifying one claim that is well-supported and one that would benefit from additional data, or (2) research and briefly describe one current policy effort to reduce microplastic pollution, then assess its likely effectiveness. Format this as a student-facing task card."

That last output needed more editing than the others — I'll get to why in a moment — but even with two rounds of adjustments, I had all four materials ready before the end of my planning period. Something that used to take me three hours took forty-five minutes, and most of that time was reading and tweaking, not generating.

Making the Three-Step Editing Pass

AI-generated differentiated materials have a particular hazard: they can feel like nobody made them for your students. They're technically correct and educationally sound and somehow completely generic. The simplified article could be for any class, anywhere. The task card could be for any teacher.

This matters. Students know when a handout came off an assembly line. The ones who already feel like they're getting "the easy version" know it most acutely.

My editing pass takes about eight minutes across all four documents, and it focuses on three things.

Names and context. I drop in a reference that anchors the material to our class. "As we discussed when we looked at the Ohio River water quality data..." That one sentence, added to the top of the scaffolded task card, makes it ours. It took me 20 seconds.

Tone. Claude's default register is slightly formal and encouraging in a way that can read as patronizing to 8th graders. I swap out phrases like "This is a great opportunity to think about..." for something a real human would say. Sometimes I just delete the sentence entirely.

One authentic example. In the extension task, I added a note referencing the Plastic Pollution Act of 2022 specifically, because we'd talked about state-level environmental legislation earlier in the semester. The AI had given me a generic prompt to "find a policy example." I made it specific. That specificity is the difference between a task that feels like homework and one that feels like a continuation of an actual conversation we were having.

Three things. Eight minutes. The materials feel like mine.

Quality Control: What to Always Read Before Printing

Here's the part I'd skip if I were skimming this chapter, so I'm putting it here where you can't miss it.

Always read the simplified version out loud. Not skimming — out loud. Claude occasionally overcorrects when simplifying and strips out so much context that the article stops making sense. In one version of the microplastics piece it removed the explanation of what a Lexile score was — which was fine, that wasn't in the original — but it also removed the explanation of what a food chain was, which left a central argument hanging. Reading aloud catches this in 90 seconds.

Check all numbers and statistics. AI can misread data presented in tables or complex sentence constructions in your source text and then reproduce that misreading in the simplified version. Numbers especially. Percentages, dates, quantities — verify them against your original.

Read every vocabulary definition. The student-friendly definitions Claude generates are almost always good. Occasionally they're confidently wrong in a way that would embarrass you. "Accumulate" defined as "to suddenly appear in large quantities" is not quite right. This happens rarely, but it happens.

Flag the sentence frames for cultural fit. This is less about accuracy than about whether the frames match how your students actually talk and write. A sentence frame that works beautifully for one population can feel stiff and foreign to another. You know your students. Scan for anything that sounds like it was written by someone who has never met a 13-year-old.

The whole quality-control pass should take you five to seven minutes if the AI did its job well. Think of it the way you'd think of proofreading an email a capable colleague drafted for you. You're not rewriting it. You're making sure it's right before it goes out under your name.

The One Thing AI Still Can't Do

I want to be straight with you here, because I've seen other teachers oversell this.

AI cannot differentiate for the student. It can differentiate the text. Those are not the same thing.

When I talk about differentiation that actually reaches kids, especially kids with learning disabilities, the most powerful version isn't about Lexile level. It's about knowing that Marcus shuts down when a text feels long, so I cut it into thirds and give him one section at a time with a sticky note that says "start here." It's knowing that Yolanda understands the content perfectly when she can discuss it first but produces almost nothing when I ask her to read silently and then write. It's knowing that Diego's IEP accommodation for extended time matters less than the fact that he needs the directions repeated privately, without an audience.

None of that lives in a text document Claude can adjust. It lives in your three years of watching these particular kids learn.

What AI produces is the raw material — the leveled text, the frames, the task card. What makes differentiation work is your judgment about which student gets which version, how you hand it out without making anyone feel sorted, when you sit down next to the student who got the simplified article and say "this one's shorter so you can go deeper" instead of "this one's easier." That's teaching. Claude cannot do it.

I also want to say something about ELL differentiation specifically. Sentence frames are a genuine scaffold. They're also a starting point, not a finishing line. The research on language acquisition — Krashen's comprehensible input hypothesis, the work coming out of the Understanding Language project at Stanford — consistently shows that ELL students need rich language exposure at and slightly above their current level, combined with meaningful interaction. A sentence frame helps a student produce a written response. It doesn't replace the conversation, the think-pair-share, the verbal rehearsal before writing. Use the frames.

Don't stop there.

The Practical Version of This

Start with one text you're already planning to teach. One. Don't try to differentiate your entire unit in one sitting.

Open Claude. Paste the article. Run the simplified version prompt first, because that's usually the most time-consuming tier to produce manually. Read it out loud. Make your three edits. Print it.

Next planning period, add the task card. The period after that, the sentence frames.

By the end of your second week using this workflow, you'll have a rhythm. By the end of a month, differentiated materials will stop feeling like a separate project that lives in that guilty desktop folder and start feeling like a twenty-minute step at the end of your normal lesson prep.

That folder on my desktop? I deleted it last spring. I didn't need it anymore. Everything I was trying to build in there — in fits and starts across three years — I can now produce in the time it takes to eat lunch.

That's not a miracle. It's just a better use of forty-five minutes.

CHAPTER V

Feedback on Student Writing — What AI Does Well, What It Doesn't

The stack of argument essays had been sitting on my kitchen table since Thursday.

Twenty-seven of them. My eighth graders had spent two weeks on a unit about whether cities should ban single-use plastics, and now their evidence was in, their conclusions were drawn, and I had until Monday to give them something useful enough to justify the assignment. Not a letter grade. Feedback. The kind that actually reaches a kid on a Tuesday morning and makes them think differently the next time they draft a paragraph.

It was Saturday at 10 p.m. My youngest was finally asleep. I had roughly ninety minutes of functional brain left.

I picked up the first essay, read three paragraphs, wrote four comments, checked my phone, read two more paragraphs, wrote two more comments. At that rate I was finishing at 3 a.m. I put down the essay. I opened Claude.

I want to be careful about how I describe what happened next, because the story has a version that sounds like a sales pitch and a version that is actually true. The actually-true version is this: AI feedback on student writing is genuinely useful for some things and genuinely harmful for others, and the line between them is sharper than most AI enthusiasts will tell you. This chapter is about that line.

What AI Gets Right, and Why It Matters

Start with the honest wins, because they are real.

AI is fast and consistent at flagging the mechanical problems that slow a piece of writing down — the vague thesis that promises an argument without making one, the body paragraph that drops a quote without any signal phrase or citation, the conclusion that restates the introduction word for word. These are the problems that appear in seventeen of your twenty-seven essays in slightly different forms, and they are also the problems that require the most repetitive written commentary from you. You have written "Your thesis needs to take a clear position" so many times that the phrase has lost all meaning, including to you.

Run those same twenty-seven essays through Claude with a well-constructed prompt and it will catch those structural issues in every single paper. Every one. No fatigue. No skipping the essay from the kid who made you smile at lunch. No subconscious leniency for the student who already seems discouraged.

Educational researcher John Hattie, whose 2009 synthesis Visible Learning pulled together findings from roughly 800 meta-analyses, placed feedback near the top of all classroom influences — effect size around 0.70, which in educational terms is enormous. But his follow-up work noted that feedback only works at that level when it is specific, timely, and focused on the task rather than the student. Those three conditions are exactly where AI shines mechanically. It is always specific ("Your thesis in sentence two states that plastic bans 'affect people' but does not specify how or argue for a position"). It is as timely as you can make it. And it is entirely task-focused — AI has no memory of the student's bad week or good attitude, which is sometimes a limitation but is sometimes a feature.

Here is what I have found AI reliably catches, across dozens of essay batches I have run through it over the past year and a half:

Thesis clarity. Claude will tell a student, in plain language, whether their thesis takes a position or just describes a topic. It gets this right probably 90% of the time.

Evidence integration. A floating quote — one dropped into a paragraph without introduction, context, or explanation — almost never slips past a well-prompted AI review.

Citation format errors. MLA in-text citations, works cited entries, basic APA. If you tell Claude which format you are using, it will spot missing page numbers and inconsistent author-name formatting faster than you can.

Repetitive sentence structure. Five sentences in a row that start with "This shows that." AI catches it. You catch it too, but you catch it at essay seven; AI catches it at essay one and essay twenty-seven with equal attention.

Topic sentence alignment. Whether a body paragraph's opening sentence actually matches what the paragraph argues. A surprising number of student paragraphs start with a claim about one thing and spend three sentences on something adjacent.

These are real problems that real students have. Catching them consistently matters.

What AI Gets Wrong, and Why It Can Hurt

Here is where I need you to slow down with me for a minute, because this part of the conversation gets skipped in a lot of the enthusiasm around AI grading tools.

AI misreads student voice. Often. In ways that can do actual damage.

I had a student — I'll call her Maya — who wrote argument essays the way some people talk: fast, direct, with a rhythm that was completely her own. One paragraph in her plastic ban essay ended like this: "So yeah, the ocean is full of straws. That's not an opinion anymore. That's just Tuesday." The structural elements were there — she had a clear claim, she had cited evidence two sentences earlier, she had a concluding pivot. But the voice was entirely hers, and it was deliberate.

Claude flagged it as "informal register inconsistent with academic argument writing" and suggested she revise for "appropriate formality."

That feedback, handed to a fourteen-year-old without context, would have told Maya that the most interesting thing about her writing was wrong. It would have smoothed her into the same beige academic voice every teacher since fourth grade has tried to enforce. And she would have gotten worse, not better.

This is not a glitch. It is a structural feature of how large language models learn to evaluate writing. They train on vast amounts of text and develop a statistical center of gravity — a sense of what argument writing typically looks like. Anything at the edges of that distribution gets flagged, whether the edge represents a problem or a strength. Cultural idioms, code-switching, deliberate fragments used for rhetorical effect, AAVE syntax employed on purpose — AI tools, including the best ones available in early 2026, treat these as errors more often than not.

I have tested this with multiple tools. The pattern holds.

There is also the problem of misread intent. A student using a rhetorical question to open a paragraph — "But who actually bears the cost?" — may get flagged for "lacking a clear topic sentence." A student who writes a short, punchy paragraph intentionally, to create contrast after two dense evidential paragraphs, may be told their paragraph is "underdeveloped." The AI cannot read the essay as a whole piece with a shape and a strategy. It reads each component against a template.

None of this means you should not use AI for writing feedback. It means you should use it carefully, for specific things, and always with your own eyes on the output before it reaches a student.

Three Paragraphs, Side by Side

Let me show you what I mean with some real comparisons. These paragraphs are from my students' essays, with all identifying details changed.

Student 1 — Call him Darius. His paragraph argues that plastic bans hurt low-income families:

"Many families who are struggling financially cannot afford reusable bags or stainless steel water bottles. A study from the University of Sydney found that reusable alternatives cost significantly more upfront than single-use options. For people living paycheck to paycheck, this cost matters. Banning plastics without addressing affordability is not environmental justice — it is environmental privilege."

AI feedback (Claude, standard argument essay prompt): "Strong use of evidence. The citation could be strengthened by including a specific statistic from the study. The concluding sentence effectively introduces the concept of environmental justice, though you may want to define this term for readers unfamiliar with it."

My handwritten comment: "Darius — that last sentence. Environmental privilege. You invented that construction right here in this paragraph, and it is doing serious rhetorical work. I want you to think about

whether you want to define it or let it hit without definition. Both are valid choices. What effect are you going for?"

The AI comment is not wrong. The page number was missing, and defining terms for an audience is a real skill. But it missed entirely the most important thing in the paragraph: that a student came up with a phrase worth keeping. My comment was about the meaning. The AI's was about the mechanics.

Student 2 — Call her Priya. Her paragraph is weaker:

"Plastic is bad for the environment. Scientists have been saying this for years. There are many reasons why banning it would help. The ocean would be cleaner and animals would not get hurt as much. This is why the ban should happen."

AI feedback: "This paragraph lacks a specific topic sentence that states a clear claim. The evidence is vague — consider citing a specific study or data point. The concluding sentence restates the topic but does not reinforce an argument. Revise to include one specific piece of evidence with a citation."

My handwritten comment: "Priya — you know more than this paragraph shows. In class you told me about the footage of sea turtles you watched. Put that here. What specifically did you see? What study did the researchers cite? Your knowledge is the material — this draft just hasn't used it yet."

Here the AI feedback is actually more complete than mine on a structural level. It caught three real problems with specificity. But it couldn't tell Priya that she had the knowledge and was holding it back, which is the thing she needed to hear. Without that piece, she might revise the paragraph and still produce something thin, because she wouldn't understand that the problem was access to her own thinking, not just a missing citation.

Student 3 — Call him Tomás. His paragraph opens with code-switching:

"My abuela always said, waste not, want not. But here's the thing — she also used plastic bags to cover her plants on cold nights, to wrap the food she brought to church potlucks, to bring groceries home on the bus. She was not wasteful. She was resourceful with what was available. Banning plastics in a neighborhood where the nearest grocery store is three miles away and nobody has a car is not the same as banning plastics in a city where Whole Foods delivers."

AI feedback: "The anecdote about your grandmother is engaging but may not be appropriate for a formal academic argument. Consider replacing the personal story with a data-driven example. The phrase 'here's the thing' reads as too conversational for academic writing."

My handwritten comment: "Tomás — this is the most powerful paragraph in your essay. The contrast between your grandmother's resourcefulness and the suburban assumption embedded in most plastic ban proposals is exactly the kind of insight this topic needs. The voice is intentional and it works. For the final draft, add one data source about food desert geography so you're combining the personal knowledge with the

research. Don't remove the story."

The AI wanted to strip out everything that made the paragraph matter. My comment wanted to protect it and strengthen it. Those are different jobs, and they require different kinds of attention.

The Triage Model

So here is how I actually use AI feedback now, after eighteen months of experimenting with it on real student work.

I think of it as triage.

In an emergency room, triage nurses sort patients by what each one needs most urgently and what kind of care can handle it. They don't do this because some patients matter less. They do it because different kinds of help require different kinds of resources, and the goal is to get the right care to the right person as efficiently as possible.

My annotating energy is a limited resource. On a Friday night with twenty-seven essays, I have maybe forty-five minutes of genuine cognitive presence before my comments get shorter and less useful. The question is not whether to use that energy — it is where to put it.

AI handles the first pass. I give Claude a prompt that describes the assignment, the grade level, the specific rubric criteria, and three or four structural issues I am watching for in this particular batch. It returns a set of comments for each student. I review those comments before they go anywhere. Some I use. Some I cut entirely. Some I use as a starting point and rewrite in my own voice.

Then I read each essay myself — but now I am reading for the things AI cannot see. The voice that is developing. The argument that almost got there. The student who is playing it safe and needs permission to take a risk. The student who took a risk and needs to know it landed.

My handwritten comments are smaller in number than they used to be. But they are better. Because I spent forty-five minutes on what only I can do instead of spending forty-five minutes on what AI can do at least as well.

That's the triage model. AI on mechanics and structure. You on meaning, growth, and the human stuff.

Privacy Is Not Optional

Before any of this works, you need to get the privacy piece right. And I will be blunt: pasting student work into the public version of ChatGPT or Claude violates FERPA if that work contains any information that could identify the student. Name, student ID, school name in the essay itself, a teacher's name referenced in the text — any of it.

This is not a hypothetical concern. FERPA requires that schools protect personally identifiable information in student education records. Submitting that information to a third-party AI company without a signed data

processing agreement puts your school at legal risk and your students at privacy risk.

The practical solution has two parts.

First: anonymize before you paste. Change the student's name to a placeholder — "Student A" — and remove any other identifying information. This takes about thirty seconds per essay and makes the feedback almost as useful.

Second: use the tools your district has actually approved. As of early 2026, Microsoft's Copilot for Education comes with a data protection agreement for school users that addresses FERPA requirements. Google Workspace for Education includes similar protections for Gemini features when accessed through a district-managed account. Some districts have also signed agreements with specific AI feedback platforms. Ask your instructional technology coordinator what is on the approved list before you build a workflow around a tool that isn't.

If your district hasn't approved anything yet, that is worth raising. But in the meantime, anonymization gets you 80% of the benefit with zero privacy risk.

One more thing. Never use AI-generated feedback as the only feedback a student receives on a significant piece of writing. AI has no idea who that student is, what they have been working on, what they struggled through to get the essay onto the page. You do. That knowledge is not decorative. It is the whole point.

The essays are still on my kitchen table some Saturday nights. But now I get through them in an hour and a half instead of four. More importantly, the comments that actually reach my students — the ones they remember, the ones they quote back at me during revision — those are still mine. AI didn't take that away. It gave me back the time to do it right.

CHAPTER VI

Rubrics and Assessments

The rubric I used for the first five years of my teaching career had three problems. First, I'd written it in the back half of an August PD session while half-listening to a presentation about data binders. Second, the language in the "3" column said things like "writing is mostly organized and mostly clear," which told students exactly nothing. Third, when I asked my colleague Marisol — who taught the same course down the hall — to score three student papers using that rubric before a moderation meeting, we were apart by two full points on four of the nine papers. Not one point. Two.

That's not a rubric. That's a suggestion.

The rubric problem is quiet and persistent. Most teachers know their existing rubrics are fuzzy. They don't have time to fix them. Building a new one from scratch can eat ninety minutes — sourcing the standard, writing four distinct performance levels for each criterion, making sure the language is parallel, checking that a student can actually read and understand what the "4" looks like before they start writing. By the time you've done all that, you've spent your entire Thursday prep period and you still have to photocopy it in the morning.

Here's what I discovered: AI won't write your rubric for you. Not really. But it will build you a credible rough draft in about four minutes, and then two targeted follow-up prompts will tighten the language to the point where inter-rater agreement actually improves. I've been through this process enough times now that I can walk you through it almost step by step.

Start Where You Should Always Start: The Standard

Before you open ChatGPT, know your anchor. For the example we'll use in this chapter, I was building a rubric for an 8th-grade argument writing unit — specifically tied to CCSS.ELA-LITERACY.W.8.1, which asks students to write arguments supported by logical reasoning and relevant evidence, with attention to acknowledging counterclaims and maintaining a formal style.

Most rubric drafts go wrong because the teacher skips this step and prompts from memory. When you paste the actual standard language into your prompt, the AI has a precise target. Without it, you get generic academic-sounding criteria that could apply to a 5th-grade opinion piece or a 12th-grade synthesis essay. The standard is the anchor. Use it.

Here is the exact prompt I used the first time I tried this:

Prompt 1 — The Draft Rubric

You are an experienced curriculum designer. Create a four-point analytic rubric for an 8th-grade argument essay. Base the rubric on this standard: CCSS.ELA-LITERACY.W.8.1 — "Write arguments to support claims with clear reasons and relevant evidence. Introduce claim(s), acknowledge and distinguish the claim(s) from alternate or opposing claims, and organize the reasons and evidence logically. Support claim(s) with logical reasoning and relevant evidence, using accurate, credible sources and demonstrating an understanding of the topic or text. Use words, phrases, and clauses to create cohesion and clarify the relationships among claim(s), counterclaims, reasons, and evidence. Establish and maintain a formal style. Provide a concluding statement or section that follows from and supports the argument presented."

Include four criteria: (1) Claim and Counterclaim, (2) Evidence and Reasoning, (3) Organization and Cohesion, (4) Style and Conventions. For each criterion, write descriptors for scores 4, 3, 2, and 1. Make descriptors observable and specific — avoid phrases like "some," "a little," or "mostly." Format as a table.

What came back was not perfect. But it was good. The "4" descriptor for Evidence and Reasoning read: "Student integrates at least two credible, relevant pieces of evidence per claim, explicitly connects each piece to the central argument using reasoning, and accurately represents sources without distortion." That's usable. That's specific enough that Marisol and I would likely land within a point of each other.

The "2" descriptor for the same criterion, though, read: "Student includes some evidence but reasoning is limited or unclear in places." That phrase — "in places" — is doing no work. It's the same vagueness my old rubric had. This is normal. First-draft AI rubric language tends to get specific at the top and mushy in the middle.

That's what the second prompt is for.

The Refinement Pass: Two Prompts, Eight Minutes

This is the part that actually takes your rubric from "probably fine" to "actually usable."

Prompt 2 — Sharpen the Middle Descriptors

Look at the score 2 and score 3 descriptors for each criterion. Right now several of them use vague language like "some," "limited," "mostly," and "in places." Rewrite every score 2 and score 3 descriptor so that the language is observable and countable where possible. For example, instead of "student includes some evidence," write "student includes one piece of evidence per claim but does not connect it explicitly to the argument." Keep score 4 and score 1 descriptors as they are.

This prompt does one thing: it forces specificity at the performance levels where most scoring disagreement actually happens. Students who earn 4s and 1s are usually easy to identify. The argument — and the grade-appeal conversation — happens in the 2-to-3 band. Language like "one piece of evidence per claim but does not connect it explicitly" gives both the teacher and the student a clear picture. That's what you're after.

After Prompt 2, you'll have a rubric that's substantially better. But there's one more pass worth making, and it's the one most teachers skip.

Prompt 3 — Student-Facing Language

Now rewrite the descriptors in the score 3 and score 4 columns only, but this time write them in second person, as if you are speaking directly to the student ("Your claim is clearly stated and distinguishes your position from at least one opposing view"). The goal is that a student can read their score and immediately understand what they did well and what the next step looks like. Keep the teacher-facing version and add a student-facing version in a separate column.

You now have a rubric with two registers: one for scoring, one for returning to students. Most teachers don't have a student-facing version because writing it twice takes forever. Here it takes about ninety seconds.

The full process — Prompt 1 through Prompt 3 plus your own editing pass — runs about twenty-five minutes the first time you do it, and closer to fifteen once you're used to reviewing AI output quickly. My old method was ninety minutes and produced worse language. I've shown this workflow to teachers in three different buildings since the fall of 2024, and the consistent report is that rubric-building drops from the "I'll do it later" pile to the "I can knock this out this afternoon" pile. That's not nothing.

One Thing You Must Always Check

Before you print anything: read every descriptor in the "1" column out loud.

AI-generated rubrics reliably produce "1" language that is either so bleak it's demoralizing ("student demonstrates no understanding of argument") or so vague it's meaningless ("student's essay lacks development"). Neither is helpful. The "1" column should describe a real, recognizable early attempt — something a student actually submitted — not a hypothetical failure. I rewrite this column by hand more often than any other, usually in under five minutes, and it's worth the time.

Also check that your criteria are genuinely independent of each other. If your "Organization" criterion bleeds into your "Style" criterion, scorers will double-penalize students for the same issue. Ask yourself: can a student score a 4 on Organization and a 2 on Style? If yes, the criteria are distinct. If the answer feels impossible, you've got overlap to fix.

Quiz and Test Generation: Bloom's Taxonomy as Your Lever

Rubrics are one side of the assessment coin. Quizzes and tests are the other, and they have a different problem: generating enough varied questions takes time teachers don't have, so the default becomes whatever's in the textbook's test bank — which is often heavily weighted toward recall and recognition, sitting at the bottom of Bloom's Taxonomy.

Bloom's six levels — remember, understand, apply, analyze, evaluate, create — are not just an academic framework. They're a practical targeting tool. When you prompt AI for quiz questions and specify the Bloom's level, you get meaningfully different output. Here's what this looks like in practice:

Quiz Generation Prompt

Generate 15 multiple-choice questions about the causes and consequences of the Missouri Compromise of 1820, appropriate for 8th-grade U.S. History students. Distribute the questions as follows: 4 questions at the "Remember" level (identifying key facts, dates, figures), 5 questions at the "Understand" level (explaining significance or summarizing events), 4 questions at the "Analyze" level (comparing perspectives, identifying cause-effect relationships), and 2 questions at the "Evaluate" level (assessing historical decisions or arguments). For each question, provide four answer choices. Mark the correct answer. After each question, note the Bloom's level.

That distribution — 4 Remember, 5 Understand, 4 Analyze, 2 Evaluate — is not arbitrary. It mirrors what a reasonably rigorous grade-level assessment actually looks like: enough recall to establish baseline knowledge, a substantial core of comprehension, and a handful of higher-order questions that stretch thinking without turning the quiz into a punishment. Adjust those numbers for your context. A pre-unit diagnostic might skew heavily toward Remember. A unit final might push more toward Analyze and Evaluate.

What you'll get back is a genuine draft test. The questions at the Remember and Understand levels tend to be solid. The higher-order questions are hit-or-miss, and the distractors — the wrong answer choices — are where things can quietly fall apart.

The Distractor Problem: The Audit Step AI Always Misses

This is the most important practical warning in this chapter.

AI-generated multiple-choice distractors have a persistent flaw: they are often either obviously wrong or accidentally correct. Both problems will undermine your assessment.

An obviously wrong distractor looks like this:

Which of the following best describes the purpose of the Missouri Compromise? A. To balance the number of free and slave states B. To declare war on Great Britain C. To establish the Federal Reserve D. To abolish slavery nationally

Option B and C are so far off-topic that any student who showed up to class will eliminate them instantly, turning your four-choice question into a coin flip. That doesn't measure what you think it's measuring.

An accidentally correct distractor is harder to spot and more dangerous. You need to read every distractor and ask: is there a reasonable argument — even a stretch of an argument — that this answer could be defended? If yes, you have an ambiguous question. Students who genuinely understood the material may argue the point, reasonably, and you'll spend fifteen minutes of test-return time in a debate you didn't plan for.

My fast audit process:

- Read each question stem without the answer choices. Predict the right answer yourself before looking.
- Read all four choices. If any distractor gives you pause, flag it.
- For every flagged item, ask one question: Could a well-prepared student defend this as correct?
- If yes, rewrite the distractor or cut the question.

This audit takes me about one minute per question for a standard quiz. A 15-question quiz: fifteen minutes of checking. Not zero, but far less than writing the quiz from scratch, and the questions you're starting with are better than what the textbook test bank usually provides.

One more thing: always run a Bloom's-level spot-check. Ask yourself whether the "Analyze" questions you received actually require analysis or whether they're dressed-up recall. I've gotten questions labeled "Analyze" that amount to "According to the text, what year did X happen?" — a Remember-level question in a coat and tie. The label is not the level. Read the actual cognitive demand.

The Prompt Library

Below are ready-to-copy prompts for the four assessment tasks teachers build most often. Paste these directly into ChatGPT or Claude, replacing the bracketed text with your specifics.

Four-Point Analytic Rubric

You are an experienced curriculum designer. Create a four-point analytic rubric for [assignment type] for [grade level] students, aligned to this standard: [paste standard]. Include these criteria: [list 3–5 criteria]. For each criterion, write descriptors for scores 4, 3, 2, and 1. Make every descriptor observable and specific — no words like "some," "mostly," or "a little." Format as a table. After the rubric, add a student-facing version of the score 4 descriptor for each criterion, written in second person.

Multiple-Choice Quiz with Bloom's Distribution

Generate [number] multiple-choice questions about [topic], appropriate for [grade level] students in [subject]. Distribute questions: [X] at the Remember level, [X] at the Understand level, [X] at the Apply or Analyze level, [X] at the Evaluate level. For each question, provide four answer choices with one clearly correct answer. Mark the correct answer and note the Bloom's level. Distractors should be plausible but clearly incorrect to a well-prepared student.

Discussion Protocol

Create a structured discussion protocol for [grade level] students to discuss [text, topic, or essential question]. Include: an opening prompt students respond to in writing for 3 minutes before discussion begins, 5 discussion questions ranging from text-dependent to evaluative, a sentence frame bank for students who need support entering the conversation, and a closing reflection prompt. Format the protocol as a one-page student handout.

Project Checklist

Create a student self-assessment checklist for [project name] in [grade level] [subject]. The project requires [brief description of what students produce]. The checklist should have 12–15 items students can check off as they complete the project. Group items into three phases: Planning, Creating, and Finalizing. Write each item as a clear, specific action a student can verify ("I have included at least two credible sources with full citations" rather than "I cited my sources"). At the bottom, include a two-question self-reflection students complete before submitting.

Exit Ticket Generator

Generate 5 exit ticket options for a lesson on [topic] for [grade level] students. Each exit ticket should take no more than 3 minutes to complete and should require students to produce original thinking, not just circle an answer. Vary the format: include at least one "explain in your own words," one "apply this concept to a new example," and one "what question do you still have?" style prompt. Note the Bloom's level for each.

Rubric Language Sharpener (for existing rubrics)

Here is a rubric I already have: [paste rubric]. The language in the score 2 and score 3 descriptors is too vague — it uses words like "some," "mostly," and "partially." Rewrite those descriptors using specific, observable, countable language where possible. Do not change the score 4 or score 1 descriptors. Return the full revised

rubric in table format.

A few quick notes on using these prompts.

The Rubric Language Sharpener is the one I give to teachers who aren't ready to rebuild from scratch. Paste in a rubric you've used for years, run this prompt, and you'll often get language improvements that would have taken you an hour to produce yourself. It's a low-stakes entry point.

The Discussion Protocol prompt works remarkably well with Claude specifically. Something about the longer, more nuanced format plays to Claude's strengths in keeping consistent register and complexity across a multi-part document. Try both tools and see what you prefer — but if you've only used ChatGPT for this, run the same prompt in Claude once and compare.

The Exit Ticket Generator is one I added after a conversation with a 6th-grade science teacher named Darnell, who said his biggest daily time-sink wasn't lesson planning, it was the small generative tasks: the exit ticket, the bellringer, the quick formative check he needed for Friday's class and was writing at 10 p.m. Thursday. Five exit tickets in thirty seconds. That's the kind of small-scale time save that adds up across a week into something you can actually feel.

What Rubrics and Assessments Actually Tell Us

There's a version of this conversation that gets overly optimistic: AI builds your rubric, AI generates your quiz, you just print and go. That's not the reality, and I wouldn't want you to show up to your first attempt expecting it to be.

The reality is that AI saves you from the blank-page problem and the parallel-language problem and the "I've been staring at this descriptor for ten minutes and I still don't know if it's clear" problem. It does not save you from the professional judgment problem. You still have to know whether the criteria you've chosen actually capture what matters in this assignment. You still have to know whether the question you're asking will give you useful information about what your students understand. The tool accelerates the drafting. The thinking is still yours.

That said — after six years of building rubrics the slow, hard way — I'll take twenty-five minutes over ninety every time.

CHAPTER VII

Parent Communication

The emails don't feel like work. That's the trap.

They feel like small things — a quick note to Marcus's dad about the missing assignments, a reply to the parent who cc'd the principal about the grade she thinks is unfair, a bi-weekly newsletter you promised at Back-to-School Night and now owe every three weeks for the next thirty-six of them. Each one takes maybe ten minutes. Except it never takes just ten minutes. You open the blank email and stare at it. You write three sentences, delete two, wonder if "I wanted to reach out" sounds condescending, settle on "I'm writing today to" which sounds corporate, delete that, and eventually produce something that gets the job done but doesn't quite sound like you.

I tracked my parent communication time for two weeks in October of my fifth year. Not emails only — I included writing the newsletter, prepping for difficult phone calls, drafting the concern note I kept putting off sending, and the follow-up messages I wrote when parents didn't respond. The total across two weeks was a little over nine hours. That's more than a full prep period per day. Gone. To writing things I would immediately forget.

The moment I handed this task to Claude, something shifted. Not because AI writes perfect parent emails — it doesn't, and I'll get to that. But because it demolishes the blank-page problem completely. You give it your messy, honest bullet points and it hands you a draft. A real, usable draft. And editing a draft is ten times faster than drafting from scratch. Always.

The Blank-Page Problem Is the Whole Problem

There's a reason parent communication keeps getting pushed to 9 p.m. on a Tuesday. It's not that the tasks themselves are complicated. It's that they require a particular kind of emotional calibration that's exhausting to summon after a full day of teaching thirty-three thirteen-year-olds.

The newsletter needs to sound warm but not chatty. The concern email needs to be serious but not alarming. The "your kid is doing great" note needs to feel personal, not like a form letter with a name swapped in. You are managing tone the way a diplomat manages bilateral negotiations, and you're doing it for thirty-two families simultaneously, each with their own history and their own tendency to read into everything.

AI doesn't share your exhaustion. It doesn't hesitate. You describe what you want — the emotional temperature, the context, the key facts — and it produces something competent and calm. You bring the judgment. It brings the first draft.

Here's exactly how I set this up.

The System Prompt That Does Half the Work

Before I paste anything into Claude, I have a standing context block saved in a note on my phone. When I open a new conversation, I paste it first. Every time. Four seconds. It makes every single output dramatically more useful than sending a cold prompt with no context.

Mine looks like this:

> *You are helping an 8th-grade ELA teacher in a Columbus, Ohio public school write parent communication. The school serves a diverse student body; many families speak Spanish or Somali as a first language, and some prefer formal written communication while others are more comfortable with a conversational tone. My name is Ms. Chen. I teach three sections of English Language Arts. When writing parent emails, use a warm but professional tone unless I specify otherwise. Assume parents are supportive and engaged unless I tell you something that suggests otherwise. Do not use educational jargon — no "scaffolding," no "metacognitive," no "growth mindset" — unless I explicitly ask. Aim for a sixth-grade reading level in all parent letters.*

That last line about reading level is not an insult to parents. It's a courtesy. Dense educational prose creates distance. Clear, plain prose builds trust. A family who feels talked at is a family who stops reading your emails. Sixth-grade reading level keeps the channel open.

With that context loaded, every subsequent prompt I write is building on a foundation Claude already understands. I don't re-explain who I am or what kind of school I'm in. I just describe the task.

Three Registers, Three Templates

Parent communication almost always falls into one of three emotional registers. I think of them as neutral-informational, warm-but-serious, and affirming. Each one needs a different approach — both in how you write the prompt and in how you edit what comes back.

Neutral-Informational: The Newsletter

I send a classroom update every three weeks to all families. Before AI, this took me somewhere between 45 and 70 minutes depending on how much I dreaded it that particular evening. Now it takes about 15 minutes total, and most of that is me deciding what to actually include, not the writing itself.

The prompt I use:

> *Write a bi-weekly classroom newsletter for 8th-grade ELA families. Here are my bullet points for what's happening: — We just finished our argument essay unit; papers were returned last Friday — Starting persuasive speech unit Monday; students will need to choose a topic by Wednesday — One mandatory library visit scheduled for next Thursday, 4th period — Reminder: late work policy*

> *resets at the semester; nothing from Q1 can be submitted after Nov 15 — Shoutout to the Period 6 class for the best Socratic seminar discussion of the year Tone: warm, practical, informative. Length: 200–250 words. Use paragraph form, not bullet points. End with a sentence that invites families to reach out with questions.*

Claude produces a clean, readable draft in about four seconds. I read it once, change anything that doesn't sound like me, add a specific student moment I remember from the week if I have one, and it's done.

The thing that genuinely surprised me: the output consistently structures information better than my from-scratch drafts did. I tend to bury the time-sensitive details halfway down. Claude leads with them. Sometimes the robot has better editorial instincts than I do at 9 p.m. on a Tuesday.

Warm-But-Serious: The Concern Email

This is the hardest one. You have a student — call him Darius — who has missed eight of the last twelve homework assignments, scored a 54 on his most recent test, and seems increasingly disconnected in class. You need to tell his parents. And you need to do it in a way that feels like an invitation to partnership, not an accusation, not a paper trail.

My raw notes before I prompt:

> *Darius, 8th grade, Period 4. 8 missing homework assignments in past 3 weeks. Test score dropped from 81 to 54 between unit 1 and unit 2. In class he's quiet — not disruptive but not present the way he was in September. I like this kid and I'm genuinely worried about him. I don't know what's happening at home. I want his parents to know I'm noticing and that I want to work together.*

The prompt:

> *Using the context from my system prompt, write a parent email based on these observational notes. Tone: warm-but-serious. This should feel like a teacher who genuinely cares about the student, not like a disciplinary notice. Do not start with "I wanted to reach out" or "I hope this email finds you well." Do not list the missing assignments like a ledger — the goal is to open a conversation, not to deliver a verdict. Keep it under 200 words.*

What comes back is usually close to what I'd write myself on a good day, when I had the emotional bandwidth and the time. It names the pattern without weaponizing the details. It extends an invitation rather than issuing a citation.

I still edit it. I always add one specific, true detail about Darius — something I noticed in class, a comment he made that told me he's still engaged somewhere underneath the quiet. That detail can't come from AI because AI has never met Darius. But I can add it in 45 seconds, and the email is meaningfully better for it.

One thing I want to say clearly: never paste a student's full name, student ID number, or any combination of information that would uniquely identify them into a general-use AI interface. I use "D." or a first-name-only placeholder. This is not paranoia — it's basic professional hygiene, the same way you wouldn't discuss a

student's IEP in the hallway at pickup. Claude's consumer interface, as of early 2026, does not use your inputs to train its models by default, but the safest habit is to keep identifying details out of it entirely. Chapter 8 goes deeper on FERPA-compliant alternatives if your district requires them for documentation tasks.

Affirming: The Good-News Note

Teachers send a lot of concern notes. We send far fewer "your kid is wonderful" notes. Parents notice.

Joyce Epstein's decades of research on school-family partnership at Johns Hopkins consistently found that positive contact from teachers — not just contact around problems — is one of the most reliable predictors of family engagement. And families who feel respected and engaged before something goes wrong are dramatically more collaborative when something eventually does.

I try to send three to five affirming notes per week. Before AI, this barely happened. Now it does. Consistently.

The prompt:

> *Write a short, personal-sounding email to the parent of an 8th-grade student named Amara. Tone: genuinely affirming, not over-the-top. She read aloud in class today and the room went quiet — other kids were actually listening. She's been shy about sharing her writing and today she took a real risk. I want her mother to know. Keep it to 100 words or less. It should feel specific, not like a form letter.*

Claude writes something warm and grounded based on the details I gave it. I read it, confirm it sounds true to Amara's actual moment, hit send. Elapsed time: maybe three minutes.

Her mom wrote back within an hour. "You have no idea how much this means to us."

Three minutes. The return is incalculable.

The Three-Minute Edit Pass

Every draft gets an edit pass. Always. Not because Claude is unreliable — it's usually quite good — but because the goal is for the email to sound like you. Parents build a relationship with a specific teacher, not a polished generic voice.

My edit pass is fast because I'm looking for exactly three things.

Does it sound like me? If a sentence construction feels stiffer or more formal than I'd ever actually write, I swap one word or recast the sentence. Two minutes, maximum.

Is there a specific detail I can add? One concrete, true observation about the student — something AI cannot manufacture — elevates the whole email from professional-but-generic to genuine. Find it. Add it. This is the thirty-second step that closes the gap between a good draft and an email someone actually reads twice.

Is anything factually wrong? AI occasionally hallucinates minor details if your prompt is imprecise. Make sure dates are correct, names are spelled right, and any policy it mentions is actually your policy. Read it. Thirty seconds.

That's the entire pass. Three minutes, start to send.

Translation: The 80% Solution (and When That's Not Enough)

About a third of my students' families speak Spanish or Somali as a first language. English-only communication is not neutral — it is a quiet barrier. I want families to actually read what I send them, which means I need to send it in a language they can read.

I use Claude to translate. Here's my honest assessment.

Spanish: quite good. The output is grammatically correct and reads naturally at the register I've set. I've had bilingual colleagues review several drafts and the consensus is consistent — it sounds like a person wrote it, not like a machine approximation. For routine updates and newsletters, I send it as-is.

Somali: more complicated. Claude can produce Somali-language text and it's often accurate at the sentence level. But idiomatic fluency — the kind that makes a letter feel like it was written for your community, not translated from somewhere else — is much harder to verify if you don't speak Somali yourself. My school has two Somali-speaking family liaisons. For any communication that is time-sensitive, emotionally serious, or involves student concern, I send the Claude translation to one of them before it goes home. It takes 24 hours. Worth it every time.

My working rule: for routine newsletters and low-stakes updates, AI translation is fine to send. For any letter touching student performance, behavior, or wellbeing — send it for review first. The families receiving those letters deserve to read something that carries the full weight of what you meant to say, not a version that's technically accurate but emotionally flattened by the distance of translation.

The prompt for translation is straightforward:

> *Translate the following parent letter into Spanish. The families who will receive this are primarily from Mexico and Central America. Use formal register (usted, not tú). Keep all proper nouns — school name, teacher name, student name — in English. If any idiom does not translate naturally, note the issue in brackets so I can choose an alternative phrasing.*

That final instruction — asking Claude to flag potential idiom problems rather than silently substitute — is one I added after an early translation where an expression came out awkward in ways I only caught because a colleague happened to read over my shoulder. Ask it to flag problems. It will. And your letters will be better for it.

What AI Still Gets Wrong

I want to be clear-eyed about this. The whole point of this book is not to oversell what AI can do.

AI-generated parent communication sounds professional and warm, but it doesn't know your families. It doesn't know that Marcus's dad works nights and you should call before 8 a.m. It doesn't know that the last time a teacher emailed Darius's mother about grades, there was a difficult conversation that ended strained. It doesn't know that Amara's family arrived eight months ago and is carrying a lot more than your ELA class right now.

You know those things. Every email you send runs through that knowledge whether you're consciously aware of it or not.

Use AI to eliminate the blank-page problem and the mechanical drafting work. Use your own knowledge to calibrate the final product to the actual human family reading it.

There is also one category of message I never draft with AI at all: preparation for a genuinely difficult phone call. If I'm calling about a potential suspension, a suspected mental health concern, or a family situation that's surfacing in the classroom in ways I can't fully name yet, I write my own notes in my own words. Some things take longer than three minutes because they should.

Your Prompt Starter Kit

Four prompts, ready to copy and adjust:

Newsletter: Write a 200–250 word classroom newsletter for 8th-grade ELA families. Tone: warm and practical. Here are my bullet points: [paste your points]. End with an invitation for families to reach out. No educational jargon.

Concern Email: Write a parent email based on these notes: [paste your observational bullet points]. Tone: warm-but-serious. Goal is to open a conversation, not list complaints. Under 200 words. Do not begin with "I hope this email finds you well."

Affirming Note: Write a short affirming note, 100 words or less, to the parent of a student named [X]. Here's what happened: [specific observation]. Tone: genuine and specific. Should not sound like a form letter.

Translation Request: Translate the following parent letter into [Spanish/Somali]. Use formal register. Keep all proper nouns in English. Flag any idioms that may not translate well.

Nine hours across two weeks down to — I checked my most recent count — just under three. That's roughly one full prep period returned to me every single week, for the rest of the school year.

Not because AI writes great emails.

Because you write great emails, faster, when you stop starting from nothing.

CHAPTER VIII

IEP and 504 Documentation

My fifth year of teaching, I carried 22 students with active IEPs or 504 plans.

I want you to sit with that number for a second. Twenty-two. In a general education ELA classroom. Each of those students came with a legal document specifying accommodations, goals, benchmarks, and required progress updates — documents that my district expected general education teachers to contribute to with professional precision. Not approximation. Not "her reading is coming along." Precise, measurable, legally defensible language that a special education coordinator, a parent, a due-process hearing officer could read and trust.

I had no training for this. My master's degree is in literature. I can tell you exactly what Toni Morrison was doing syntactically in the first paragraph of Beloved. I cannot tell you, or I couldn't then, what differentiates a well-written present level of performance from a vague one — not in the technical sense the documents required.

So I did what most general education teachers do. I wrote slowly, anxiously, and badly. I stared at blank text fields until my eyes dried out. I copied sentence structures from old IEPs and swapped in names, which is not the same as writing accurately. And every November, when progress reports were due, I spent three evenings doing paperwork I should have been spending with my own family.

The fifth year was the year I almost quit.

Not because of the students. Never because of the students. Because of the sheer administrative weight of serving them properly.

What Changed

I want to be careful here, because this is the chapter where I have to say something I don't say lightly: AI helped me do this job better. Not faster at the expense of quality. Genuinely better. More accurate. More useful for the families reading these documents.

But I also have to say the other thing up front, before we get to any prompts: the legal and ethical guardrails around student information are not optional, and they are not a formality. FERPA — the Family Educational Rights and Privacy Act — governs what can happen to student data, including which third-party software platforms can process it and under what conditions. Violating FERPA isn't a paperwork error. It's a legal breach that can cost your district federal funding and, more importantly, exposes your students' most sensitive information to platforms that have no accountability to them or their families.

So we start there. Before we get to anything else.

The Rule You Cannot Bend

Here it is, plainly: do not paste identifiable student information into a consumer AI tool.

Consumer tools include the free tier of ChatGPT at chat.openai.com, the free tier of Claude at claude.ai, Google Gemini in its standard personal account form, and any other AI interface you're accessing through a personal account that is not covered by a data processing agreement with your school or district.

"Identifiable information" means name, date of birth, student ID number, specific diagnosis, detailed disability description, or any combination of details that could identify the student. A sentence like "Marcus, age 13, reading at a 4th-grade level, diagnosed with dyslexia and ADHD, struggles with multi-step directions" should never go into a tool your district hasn't approved.

What your district may have approved — and this varies, so check — might include Microsoft 365 Copilot for Education (Microsoft publishes a Student Data Privacy Consortium agreement covering it), Google Workspace for Education AI features used within a district-managed account, or enterprise-tier versions of Claude and ChatGPT that include signed Data Processing Agreements giving schools the contractual protections FERPA requires.

When I talked to my district's technology coordinator in 2024, she gave me a list of three approved tools. That list is what I use. If your district doesn't have a list yet, ask your tech coordinator or your union rep to start that conversation. It is their job to have an answer.

Everything in this chapter assumes you are using a district-approved tool, or that you have anonymized the student information before it goes anywhere near an AI interface. I'll show you exactly how to anonymize in a moment.

What AI Is Actually Good For Here

Let me tell you what the problem is, and then you'll see immediately why AI fits into it.

A present level of performance summary — sometimes called a PLAAFP, or Present Levels of Academic Achievement and Functional Performance — is supposed to do several specific things at once. It needs to describe where a student is right now, in measurable terms, referenced to grade-level expectations. It needs to explain how the student's disability affects their access to the general education curriculum. And it needs to be written in language a parent without a special education background can understand.

That is actually a hard technical writing problem. You need specificity, accessibility, and accuracy simultaneously. Most teachers — me included — have the observational data. We've watched this student for weeks. We know they can decode fluently but lose comprehension above a 6th-grade Lexile. We know they shut down during timed tasks. We know they excel in discussion but struggle to transfer ideas to the page. We know these things in our bones.

What we often can't do quickly is translate those bone-deep observations into the structured, formal, measurable prose the document requires. That's the gap. And it's the gap AI closes remarkably well.

The Anonymization Step

Before you write a single prompt, do this: replace every student identifier with a placeholder.

Use a system that takes you less than ninety seconds. I call my placeholder students by fictional first names I've never had in class — names far enough from my actual roster that I won't confuse them. Or I use letters. "Student A," "the student," "this learner." Replace the grade level with a generic descriptor ("middle school student"). Remove any specific diagnosis from what you paste in — you can describe functional impacts without naming the diagnosis in your prompt.

Your rough notes might look like this in their original form:

> *Marcus, 8th grade, IEP for learning disability in reading/writing. Reads at about 5th-grade level. Great in class discussion but his written work doesn't show what he knows. Gets stuck organizing ideas. Needs sentence frames. Doesn't finish timed tasks.*

What you paste into the AI tool looks like this:

> *Student is in 8th grade. Reads approximately two to three grade levels below grade level expectations. Participates actively and insightfully in verbal discussions but struggles to transfer ideas to written form. Has difficulty with organizational structure in multi-paragraph writing. Benefits from sentence frames and graphic organizers. Does not consistently complete timed writing tasks.*

Same information. No name, no diagnosis, no identifying details. The output the AI produces goes back into the document only after you've re-personalized it with the student's actual name and verified it against what you know.

Ninety seconds of anonymizing. It matters.

A Prompt That Actually Works

Here is the prompt I use when I need to turn rough observational notes into a PLAAFP-style summary. Use it verbatim, or adapt it:

"You are helping a general education teacher draft a present level of performance summary for an IEP. The summary should: describe the student's current academic performance in concrete, measurable terms; explain how the student's learning challenges affect access to grade-level curriculum; use clear language accessible to parents without special education backgrounds; and be written in third-person, professional prose. Here are my rough observational notes. Please turn them into a structured paragraph of approximately 100–150 words.

Notes: [paste your anonymized notes here]"

That's it. No extra instructions. No jargon. The specificity about length and about the parent-readability requirement does the work.

Here is a before-and-after using the anonymized Marcus example. This is close to a real exchange I had in November 2024, reconstructed with details changed.

Before (my raw notes):

> *Student is in 8th grade. Reads approximately two to three grade levels below grade level expectations. Participates actively and insightfully in verbal discussions but struggles to transfer ideas to written form. Has difficulty with organizational structure in multi-paragraph writing. Benefits from sentence frames and graphic organizers. Does not consistently complete timed writing tasks.*

After (AI draft, lightly edited):

> *This student currently reads approximately two to three grade levels below 8th-grade expectations, as demonstrated by classroom assessments and informal reading inventories. He demonstrates strong verbal reasoning skills and contributes meaningfully to class discussion, consistently offering on-grade-level analysis when ideas are explored orally. His disability affects his ability to organize and express ideas in written form: multi-paragraph writing tasks reveal difficulty with structural sequencing, and his written work does not yet reflect the comprehension he demonstrates verbally. He benefits from pre-writing supports including sentence frames and graphic organizers. Timed writing tasks present a consistent challenge and may not accurately reflect his understanding of the content.*

That draft took eleven seconds to generate. I read it carefully, checked it against my actual knowledge of the student, swapped in his name, and made two small edits — one phrase felt too clinical, one felt like it undersold a genuine strength. The final version went to the special education coordinator for her review, as it always does. She said it was one of the cleaner PLAAFP contributions she'd received from a general education teacher that cycle.

I'm not bragging. I'm showing you what the ceiling looks like when the process works.

Goal Progress Narratives

Progress notes are a different beast from PLAAFP summaries. Where the PLAAFP describes where a student is, progress notes must describe change — movement toward or away from a specific measurable goal — over a defined reporting period. They require you to reference the goal language from the IEP itself, connect your observations to that language, and say something meaningful about next steps.

Most teachers write these in a fog of vague language. "Making progress." "Continuing to work on." "Shows improvement in some areas." That language is almost useless — legally and practically. It tells the family nothing. It tells the next teacher nothing.

Here is a prompt that produces something better:

"You are helping a general education teacher write a progress note for an IEP goal. The note should: reference the specific goal (provided below); describe observable evidence of progress or lack of progress during the reporting period; avoid vague language like 'making progress' or 'showing improvement'; and conclude with one sentence about the instructional focus for the next period. Keep the note to 75–100 words. Write in professional third-person prose.

Goal: [paste the exact goal language from the IEP]

My observations during this reporting period: [paste your anonymized notes]"*

The goal language you paste in keeps the AI anchored to the actual document. Without it, AI will write something generic. With it, the output connects directly to the measurable language the team originally agreed on, which is what a progress note is supposed to do.

What AI Gets Wrong Here

I promised you honesty, so here it is.

AI produces fluent, professional language, which can make you feel more confident in a document than you should. Fluency is not accuracy. A beautifully worded progress note that misrepresents a student's actual trajectory is worse than an awkward one that tells the truth — because the beautiful one might go into a file and influence decisions without anyone catching the error.

Read every line against your own direct knowledge of the student. Ask: Is this true? Is this specific? Does this match what I've seen? If something the AI wrote is technically accurate but leaves out a complicating truth, fix it. If the AI turned your observation about a struggling student into something that reads more optimistically than the reality warrants, fix it. The document serves the student, not your desire to submit a clean draft.

AI also cannot make professional judgments about what data means. It can describe what you describe. It cannot interpret an observation it doesn't have. If you didn't write in your notes that the student's performance dropped sharply after a change in medication schedule, the AI doesn't know that. If you didn't note that the student excels in one-on-one settings but struggles in a full class, the AI doesn't know that. Your observational knowledge is the content. AI is the structure. The distinction matters enormously.

And then there is the question of what counts as compliant language in your specific district and state. IEP documentation requirements vary. What's acceptable goal-progress language in Ohio may not match the format a district in California uses. AI doesn't know your district's requirements unless you tell it — and even then, verify with your special education coordinator before anything goes into a final document.

The Workflow, Start to Finish

Here is how I run this from the beginning of a reporting period to the submitted document:

During the grading period: I keep a running notes document for each student with an IEP or 504. It takes literally two minutes per student per week — one or two observations jotted after a class where something notable happened. By the end of the reporting period, I have four to six weeks of notes per student.

One week before documentation is due: I open my approved AI tool. I anonymize each student's notes using the placeholder system. I run the appropriate prompt — PLAAFP or progress note — and save the output in a working document.

Review day: I read every AI-generated draft against my own notes and my knowledge of the student. I edit. This is not a rubber-stamp step. This takes time — less time than writing from scratch, but real time. Budget about five minutes per student for review and light editing.

Submit to special education coordinator: Nothing goes into an official IEP without my coordinator reviewing it. That's her job and her expertise. I'm a contributing team member, not the document's author in isolation. The IEP team writes the IEP.

From four evenings of anxious work to one focused afternoon. That's what the shift looked like for me.

On the 504 Side

504 plans are less complex than IEPs in terms of documentation volume — they don't require the same scope of goal language and progress narratives — but they come with their own paperwork. Accommodation logs, progress check-ins, parent communication records. The same anonymization discipline applies. The same AI-as-draft-not-final principle applies. For 504 documentation, I use a simpler prompt:

"Draft a brief progress note (50–75 words) for a student's 504 accommodation review. The student receives the following accommodations: [list accommodations]. My observations on how well the accommodations are working: [paste anonymized notes]. Note whether accommodations appear effective, partially effective, or in need of adjustment."

Short. Specific. Reviewable in ninety seconds.

The Bigger Point

I think about a parent sitting across from me at an IEP meeting, holding a progress note I wrote at midnight after three glasses of wine because I had no time and no support. I think about what they deserved versus what I had the capacity to give them in that moment.

That's the real argument for using AI here — not efficiency for its own sake, but quality in service of the students and families who depend on these documents. When I'm not drowning in the mechanics of documentation, I have more mental space to think about what I actually know about each child. The AI handles sentence-level architecture. I handle meaning.

The legal guardrails exist for good reason. The professional judgment requirement exists for good reason. Neither of those things is in tension with using AI. They're just the conditions under which you use it responsibly.

Your job as a teacher of record is to know your students well enough to verify what AI produces. If you know your students well, you'll catch the errors. If you've been too overwhelmed to pay close attention — which the documentation load itself sometimes causes — this workflow gives you a little breathing room to look up from the paperwork and back at the actual child.

That's worth protecting.

CHAPTER IX

The Hard Conversation: Students Using AI

Every August, on the first full week of school, I ask my eighth graders the same question before I've handed out a single syllabus.

"How many of you have already used ChatGPT — or any AI — to help with homework?"

The first time I asked, in the fall of 2023, I expected two or three sheepish hands. What I got was twenty-six hands out of twenty-eight kids. Not tentative, either. Confident. A couple of kids in the back were almost proud. One boy — Marcus, a relentless arguer who would go on to write the best persuasive essay I read that year — looked me dead in the eyes while his hand went up, as if daring me to make something of it.

I didn't. And that choice changed everything about how I handle this subject.

What the Hands Told Me

If you run the same informal poll in your classroom, you'll probably get similar numbers. A 2024 survey by Common Sense Media found that 53% of teens reported using AI tools for schoolwork at least once a month, with middle and high schoolers significantly outpacing elementary students. By the time a kid hits eighth grade, ChatGPT is roughly as exotic as Google Translate — which, if you remember, we also panicked about, also tried to ban, and also eventually made peace with.

The instinct most of us have when we see that forest of raised hands is fear, followed quickly by its close cousin, punishment-planning. I've been in department meetings where colleagues started drafting zero-tolerance policies before anyone had actually read a piece of student-generated AI writing. I understand the impulse. It feels like cheating. It feels like the erosion of something important. And honestly, sometimes it is cheating. I'm not going to pretend this is simple.

But. Twenty-six out of twenty-eight kids raised their hands without any guarantee I wasn't going to send them to the office. That told me something more interesting than "students cheat." It told me that the kids were operating in a world we hadn't caught up to yet, and that defaulting to punishment without a framework was going to make me an obstacle rather than a teacher.

Curiosity first. Policy second. That's the order that actually works.

Building a Policy That Can Survive Contact with Reality

Here's what doesn't work: a blanket ban with no distinctions. I've watched it play out. Teachers announce that AI use is academic dishonesty, full stop. Students who used AI for legitimate brainstorming feel vaguely criminal. Students who used it to write their entire essay feel barely inconvenienced, because detection is

unreliable (more on that in Chapter 10). The policy looks rigorous on paper and accomplishes almost nothing in practice.

What works is a policy with teeth, calibrated by task.

The framework I use sorts every assignment into one of three categories, and I'm explicit with students about which category each task falls into.

AI-closed tasks are assessments where I need to see your thinking and only your thinking. Timed in-class writing, reading responses where you need to demonstrate you understood the actual text, tests. On these, AI assistance of any kind is off the table, same as opening your phone during a quiz. The reasoning is obvious and I state it plainly: I can't assess what you know if something else is generating the answer.

AI-transparent tasks are tasks where you may use AI as a starting point or a tool, but you must document how you used it and what you changed. A research brainstorming session, generating counterarguments to pressure-test your own essay, using AI to check grammar on a final draft. On these, I require a short "AI log" — two or three sentences explaining what they asked, what it gave them, and what they did differently. This is not punitive. It's workflow documentation, the same thing professionals do.

AI-integrated tasks are assignments I've specifically designed around AI use, usually because the learning objective is about the skill of working with AI, not the product it produces. More on how to design these in a minute.

This three-part system gives students a coherent mental model instead of a foggy sense of "AI is probably bad but I'm not sure where the line is." Fog is where cheating lives. Clarity reduces it.

Making the Policy Stick: The First-Week Protocol

I spend part of one class period in the first week running what I call the AI Honest Conversation. Not a lecture. A protocol.

It goes like this.

I give everyone an index card and ask them to write down one time they've used AI in a way they were proud of — it can be outside school, it can be anything — and one time they used it in a way they weren't sure about. Anonymous. They fold the cards and drop them in a box on my desk.

I read a handful aloud, anonymized, and we talk. The proud examples are usually creative: someone used ChatGPT to help plan a birthday party, someone used it to debug their Minecraft mod code. The uncertain examples are almost always school-adjacent: "I wasn't sure how to start an essay so I asked it for an introduction." "I asked it to summarize a chapter I hadn't read."

From those uncertain examples, we build the policy together. What makes the first scenario (starter help) feel different from the second scenario (replacing the reading)? The students will tell you. They know. Eighth graders have a sharp moral intuition when they're not on the defensive, and asking them to reason through the

ethics rather than hand them a list of rules produces buy-in that a syllabus policy never will.

At the end of the period, I distribute my Classroom AI Agreement — which I'll share in full below — and we go through it line by line. They sign it. I sign it. I keep a copy. This sounds bureaucratic, but it matters: when I later have a conversation with a student about a suspicious assignment, I can start from a document we both agreed to, rather than a rule I'm inventing on the spot.

Redesigning Assignments So Shortcuts Backfire

This is the part most policy guides skip, and it's actually the most powerful lever you have.

AI produces mediocre output on tasks that require personal specificity. It cannot tell you what Marcus argued in third period on Tuesday and how you'd respond to him. It cannot describe the moment in The Outsiders that surprised you and why. It cannot write a reflection grounded in your own observation of your science lab's actual data, with the actual numbers your group collected.

When I redesigned my persuasive essay assignment in 2024, I added three required elements that AI simply cannot fake without substantial human input: a reference to a class discussion or peer's argument from our Socratic seminar, a personal anecdote that connects to the issue, and one piece of evidence from our specific classroom nonfiction packet (not widely available online). A student who dumped the prompt into ChatGPT and submitted the output unedited would fail on those three criteria automatically.

This isn't about being adversarial. It's about making the assignment actually do what assignments are supposed to do — reveal the student's thinking.

Some concrete redesign moves that work at almost any grade level:

Require process documentation. Ask for a photo of their prewriting, a voice memo explaining their reasoning, a draft with tracked changes. AI doesn't leave a paper trail. Students do.

Build in specificity anchors. Every major assignment should reference at least one thing that happened only in your classroom — a particular text you discussed, a piece of data your class generated, a position a classmate took. These are impossible to anticipate from a generic prompt.

Make the stakes of the first draft low. When students know the rough draft doesn't count for much, they have less incentive to offload it to AI and more freedom to try messy, genuinely their own ideas. This also gives you a baseline to compare against the final draft, which is useful if something shifts dramatically.

Use verbal or interactive follow-up. Not as interrogation — as routine. "Tell me about what you were thinking when you wrote this part." Make it a normal classroom move, not reserved for suspicious papers. When students know you're going to talk about their writing, they're more invested in actually understanding it.

The Conversation You Have to Have

At some point, no matter how solid your policy and assignment design are, you will read a piece of student writing that doesn't sound like that student.

Maybe the vocabulary jumped three grade levels. Maybe every paragraph starts with a claim-evidence-analysis structure that was never discussed in your class. Maybe the citations are formatted perfectly and that student has never formatted a citation correctly in their life.

Do not lead with an accusation. I cannot say this firmly enough.

"Did you use AI?" puts a student immediately on the defensive, because the only answer that helps them is "no" — and so "no" is what you'll hear, regardless of the truth. You've created a situation with no good outcomes.

Instead, try this script, which I've refined over two years of these conversations:

"I want to talk with you about this piece because I'm trying to understand your thinking. Can you walk me through how you developed your argument in the second paragraph?"

If the student wrote it, they can walk you through it. Maybe haltingly, maybe with gaps, but they'll have something — a memory of a source, an idea that changed mid-draft, a sentence they almost cut. If they didn't write it, they'll struggle. They'll give you vague answers. They'll describe the paragraph's content rather than its creation.

You're not trapping anyone. You're gathering information. And most of the time, what you'll find is more nuanced than "they cheated": they used AI to help with a structure they didn't understand, or they used it to check their grammar and accidentally overwrote their voice in the process. These are teachable moments, not tribunal moments.

If the conversation confirms that the student submitted AI output as their own work — particularly on an AI-closed assignment — then you apply your consequence consistently and without drama. The conversation already happened. The evidence is the student's own inability to explain their own work. You don't need a detection score. You need what they just told you.

Talking to Parents Before It's an Incident

The best parent conversation about AI is the one you have in September, not the one you have when their child is sitting outside the principal's office.

I send a brief paragraph home in my back-to-school letter — folded into my general class overview — that reads something like this:

"AI tools like ChatGPT are part of our students' world, and I address this directly in class. I have a classroom AI policy that students have reviewed and signed, and I'll be in touch if any concerns come up. If your child ever has questions about what's appropriate to use AI for on a specific assignment, encourage them to ask me directly — I'd rather have that conversation early than late."

That's it. Short. Not alarming. It does three things: it signals that you're aware of the landscape, it references a system that's already in place, and it positions direct communication as the normal course of action.

When you do have to contact a parent about a specific incident, you lead with the documentation. "We have a classroom AI agreement that [student's name] signed on September 5th. Here's the section that's relevant." This is not about embarrassing the student or the parent. It's about grounding the conversation in a shared agreement rather than your subjective impression that something felt off.

Most parents respond well when they understand that you've been proactive rather than reactive. The parents who push back hardest are usually the ones who are blindsided — who feel like the rules weren't clear, or that AI policy is a moving target their kid couldn't have anticipated. Clear communication in September dramatically reduces those defensive responses in March.

The Classroom AI Agreement

This is the document I actually use, lightly modified to remove identifying school information. You should adapt it to your grade level, subject, and district policy — run it by your department head before you distribute it — but this is a working starting point.

Our Classroom Agreement on AI Tools Ms. [Teacher] — 8th Grade ELA

AI tools like ChatGPT, Claude, and others are powerful, and they're going to be part of your professional life. In this class, we're going to learn to use them well — which means learning when to use them, when not to, and how to be honest about it either way.

What I agree to as your teacher:

- I will tell you clearly for every major assignment whether it is AI-closed, AI-transparent, or AI-integrated.
- I will not accuse you of AI use without having a direct conversation with you first.
- I will treat AI as a literacy issue, not just a discipline issue.

What you agree to as a student:

- I will follow the category designation for each assignment.
- On AI-transparent tasks, I will complete an AI log explaining how I used any AI tool and what I changed.
- I will not submit AI-generated text as my own original work on AI-closed assignments.
- If I'm unsure whether a particular use of AI is appropriate for a specific task, I will ask before submitting, not after.
- I understand that if I cannot explain my own work in a follow-up conversation, that is itself evidence that the work isn't mine.

Student signature: _____ Date: _____ Teacher signature: _____ Date: _____

One page. Twelve sentences. The students can actually read it.

What This Really Takes

I want to be honest about something before you close this chapter.

None of this eliminates the problem. Students will still sometimes make bad choices. Detection will still sometimes be impossible. You'll read a paper that feels wrong and never be able to prove it, and that is genuinely frustrating.

But the framework I've described does something more valuable than catching cheaters: it produces students who have actually thought about why intellectual honesty matters, rather than students who have calculated the odds of getting caught.

Marcus — the kid from that first day with his hand confidently in the air — wrote me a note at the end of eighth grade. He said the thing he learned in my class wasn't how to write a persuasive essay. It was "how to think about tools without being dumb about them." I'm not going to pretend that's entirely attributable to my AI policy. But I think the honest conversation we had on day one, the framework that treated him as someone capable of reasoning rather than someone to be surveilled, had something to do with it.

That's the bar. Not zero AI use. Not perfect compliance. Students who can think about this stuff.

That's actually a higher bar than catching cheaters. And it's worth more.

CHAPTER X

Honest Detection (vs. Accusations You Can't Prove)

Picture this: a seventh grader named Marcus submits a literary analysis essay on The Outsiders. It's good — actually, genuinely good. Better than anything he's turned in before. His teacher runs it through an AI detector. The tool returns a score: 94% AI-generated. The teacher prints the report and calls the student to the hallway.

Marcus cries. He says he wrote it himself, over two nights, because his older cousin helped him outline it. The teacher doesn't fully believe him. The conversation goes badly. His mother calls the principal. Three weeks of tension follow, and Marcus stops trying hard on writing assignments for the rest of the year.

I didn't witness that scene myself, but a version of it — same tool, same misplaced certainty, same aftermath — has played out in schools across the country since AI detectors became a product teachers could buy or get for free. I've heard stories like it from colleagues in my building, on teacher forums, in the comment sections of education newsletters. The details change. The damage is consistent.

This chapter is about what those tools are actually measuring, why the answer should worry you, and what actually works instead.

What Detectors Are Measuring (It's Not What You Think)

AI detection software doesn't read an essay the way you do. It doesn't recognize ideas, evaluate argument quality, or spot a thesis that sounds too sophisticated for a particular student. It runs statistical analysis on text patterns — specifically, two measurements that researchers call perplexity and burstiness.

Perplexity, in this context, measures how predictable the word choices are. Language models like ChatGPT are trained to generate high-probability sequences — words that are statistically likely to follow the words before them. Human writers, particularly when they're writing expressively or taking risks, tend to make less predictable word choices. Low perplexity in a text signals to the detector that the word choices look like what a model would generate. High perplexity looks more human.

Burstiness measures variation in sentence length and complexity. Human writing tends to be irregular — short sentences next to long ones, simple clauses followed by complicated nested structures. AI-generated text tends toward a more consistent cadence. Detectors flag low burstiness as a warning sign.

Here's the problem, and it's a big one: both of those signals can describe human writing that is just formal, cautious, or produced by someone writing in their second or third language. A student who has learned English through structured grammar instruction — whose native Somali, Spanish, or Mandarin trained her toward precise, orderly sentence construction — will often produce low-perplexity text because her English

vocabulary draws on a more deliberate, careful word-selection process. She's not using AI. She's just writing the way a careful non-native writer writes.

In July 2023, a team of researchers at Stanford published exactly this finding in the journal Patterns. They tested seven widely used AI detectors against a set of essays written by non-native English speakers for the TOEFL exam — genuine human writing, definitionally. The detectors flagged those essays as AI-generated at rates that reached over 60 percent for some tools. The same detectors applied to essays by native English speakers returned false positive rates in the low single digits. The tools weren't finding AI. They were finding a writing style that looks, to an algorithm, statistically similar to AI output.

That study came out two and a half years ago. The tools have been updated since then. The fundamental measurement problem has not gone away.

OpenAI Walked Away From Its Own Classifier

There's a detail that doesn't get mentioned enough in school policy discussions: OpenAI — the company that makes ChatGPT, the tool at the center of most academic integrity worries — built and released its own AI text classifier in January 2023. By July 2023, they shut it down. Their public statement said it was "no longer available" because it was "not fully reliable." They cited a low accuracy rate and specifically acknowledged the risk of false positives.

The company most responsible for the concern decided the detection problem was not solved. The third-party tools that schools are licensing and building policy around have not solved it either. They have improved their marketing. That's different.

Turnitin, which released AI detection as part of its existing plagiarism platform in April 2023, has been more transparent than most. Their published documentation acknowledges a false positive rate and recommends that instructors treat AI detection scores as a starting point for conversation, not a final determination. Read that carefully: the company that sells the tool is telling you not to use it as evidence. That's a remarkable disclaimer to bury in documentation that most teachers and administrators never read.

GPTZero, created by Princeton student Edward Tian in early 2023 and now a commercial product, has similar documentation. Originality.ai, another commonly used tool, markets aggressively to schools while its own accuracy disclosures are, to put it charitably, optimistic.

None of these are bad-faith operations. Detection is a genuinely hard computational problem, and the people building these tools are working on it. But "we're working on it" is not a sufficient foundation for accusing a child of academic dishonesty.

The Cascade Problem

The false positive issue isn't just about the single student who gets wrongly accused. It's about what accusation does to a classroom.

Trust is the substrate of everything that happens in school. When a student believes that working harder could get them accused of cheating more — that a better essay is more suspicious than a worse one — you have inverted the incentive structure of education. Some students respond by sandbagging their writing deliberately. Some respond by actually using AI to avoid looking like they used AI, which is a special kind of perverse outcome. Most just absorb a low-grade lesson about institutions: the system doesn't actually care about your effort; it cares about its metrics.

I'm not being melodramatic. I watched a version of this play out in a colleague's classroom after a department-wide policy went into effect requiring AI detection scores for any essay flagged as "suspicious." The submissions that followed were shorter, safer, blander. The students had learned something, but it wasn't anything the curriculum intended.

Why I Still Have the Tool Open Sometimes

I want to be honest here: I haven't deleted my GPTZero account. I still occasionally run a piece of work through it when something feels off — not in the student's writing quality, but in the meta-signals. A student who has never submitted more than two paragraphs suddenly turns in a fully structured, five-paragraph essay with no revision history visible in Google Docs. A piece that contains vocabulary the student has never used in any other context. Something that reads like it was written for a different teacher, a different class, maybe a different grade.

When I run text through a detector in those situations, I'm not building a case. I'm letting myself be curious. A high score doesn't become evidence. It becomes a prompt — a reason to have a conversation I was going to have anyway.

That's the appropriate use of these tools. Private. Tentative. A curiosity trigger, not a verdict.

The Authorship Conversation

The thing that actually works is simple enough that it feels almost too obvious: ask the student to talk about their own work.

I call it the authorship conversation, though a colleague of mine who runs a middle school humanities program calls it a "walk me through it" check, which is maybe more student-friendly language. The core idea is the same. If a student wrote something, they know things about it that don't appear in the text itself. They know why they chose the opening anecdote. They know what they originally thought their thesis was before it shifted. They know which sentence took the longest and why. They know what they googled when they got stuck in the second body paragraph.

A student who used AI to generate a piece and submitted it without engaging with it genuinely cannot answer those questions. Not because they're lying — though sometimes they are — but because the thinking never happened. There's nothing to retrieve.

Here's how I structure it. When I want to have an authorship conversation, I pull the student aside for about five to eight minutes. I tell them I want to hear more about their process because I found the piece interesting. That's usually true. I ask three to four questions:

— Walk me through how you got started. What was your first idea for this? — Which part of this was hardest to write? What did you do when you got stuck? — Is there anything you wanted to say that didn't make it into the final draft? — If you could change one paragraph, which one and why?

These questions have no right answers. What I'm listening for is fluency and specificity — the casual way a person talks about their own thinking when they actually did the thinking. Students who wrote their own work almost always give uneven, specific, sometimes self-deprecating answers. "The intro was bad and I kept changing it." "I wasn't sure if I was supposed to use first person." "I had this whole part about Ponyboy's relationship with his brothers and I cut it because it was getting too long." That texture is what genuine authorship sounds like.

Students who didn't write the work often go one of two directions: they give very general, smooth answers that sound like they're describing an essay they read, or they go silent and admit it quickly. Either outcome is informative.

Here's the important piece: an authorship conversation is not an interrogation, and it should never be framed as one. I go into it genuinely curious. Sometimes I discover the student is articulate about their own work in ways the essay doesn't reveal, and I come away more impressed, not suspicious. That happens. It should happen. The conversation serves the student even when there's nothing to find.

When I do find clear evidence — when a student cannot explain their thesis, doesn't know what the essay argues, or volunteers an admission — I handle it according to my school's academic integrity policy, the same way I would handle copied text. The authorship conversation is the discovery mechanism. The policy handles the consequence.

What it is not, ever, is a situation where I print out a detection score and present it as proof. That path leads to Marcus crying in a hallway, and I don't want to go there.

A Note on Written Authorship Checks

If you teach large classes or have limited time for individual conversations, a version of this works in writing too. In-class writing prompts that ask students to annotate their own submitted work — "circle the sentence you're least confident about and explain why" — give you something to compare against the submitted piece. Students who genuinely engaged with their own draft can do this. Students who copy-pasted something at eleven o'clock the night before generally cannot, because they don't know which sentence they're least confident about. They don't know anything about the sentences.

You can also build low-stakes process documentation into the assignment itself: a brief writer's memo submitted with the final draft, a required revision note explaining one change made between drafts, a

photograph of a handwritten outline. None of these are foolproof — a sufficiently motivated student can fake process documentation too — but they raise the cost of shortcuts significantly, and they're infinitely more defensible than a detection percentage.

For Administrators and Department Heads

If you've read this far and you're an administrator rather than a classroom teacher, what follows is written specifically for you.

Policies that require AI detection scores as part of academic integrity proceedings are policies built on a measurement tool that their own creators describe as unreliable. If your district's legal team reviewed your plagiarism policy before adopting it, ask whether they reviewed the published accuracy documentation for the specific detection tool you're using. Ask whether the policy accounts for differential false positive rates across student demographics. Ask what the appeals process looks like for a student who receives a 94% AI score on work they genuinely wrote.

These are not hostile questions. They are exactly the questions a fair policy requires.

Teachers in your building who want to push back on AI detector mandates can use the following framework: detection scores are one data point, not a determination; any integrity proceeding that rests on a detection score without an authorship conversation lacks procedural fairness; and the Liang et al. (2023) research in Patterns establishes documented demographic bias in the tools currently on the market. That research is peer-reviewed, publicly available, and directly relevant to how these tools perform in diverse classrooms.

There are also districts that have handled this thoughtfully. Some have moved to a model where detection output requires a follow-up conversation before any process begins — exactly the triage model I described above. That's a policy defensible in front of a parent, a school board, and a lawyer. "The algorithm said so" is not.

What This Means for Your Daily Practice

A few practical shifts that come out of everything above:

Build authorship conversations into your regular routine, not just your suspicious moments. When you check in verbally with students about their writing regularly, the conversation stops feeling like an accusation and starts feeling like class. Students who write their own work appreciate it. Students who don't get used to the expectation that they'll be asked.

Design assignments with process embedded in them. Due dates for outlines, rough drafts, and revisions create a paper trail that makes wholesale AI substitution much harder. This is covered in more depth in Chapter Nine, but the overlap is real: assignment design and integrity practice are the same conversation.

When something feels off, be curious before you're certain. Five minutes of genuine conversation will tell you more than any tool. And if you're wrong — if the conversation reveals a student who worked hard and

wrote something better than you expected — you've just given them the experience of being taken seriously. That's not a bad outcome.

The goal of all of this is not to catch students. It's to teach them. Sometimes those things overlap. Usually, if you've done the design work and built the relationships, they don't have to.

One last thing. When Marcus — or a student like him — sits across from you, the most important question you can ask is not "did AI write this?" The most important question is "tell me what you were thinking when you wrote this." The answer will always be more useful than a percentage score. And it treats the student the way students deserve to be treated: as the presumptive author of their own thinking, until proven otherwise by something more reliable than an algorithm that four researchers in California already showed has a demographic problem.

That's not a low bar. It's the right one.

CHAPTER XI

Teaching ABOUT AI in Your Subject

Last October, a student named Marcus raised his hand in the middle of my persuasive writing unit and asked, with genuine curiosity and zero shame, whether he could use ChatGPT to help him "think of reasons" for his argument essay. I told him we'd talk about it. Then I spent the next three minutes thinking about how to answer him and realized I had no idea what I was going to say.

That pause — that three-minute blank — is where this chapter lives.

Because Marcus wasn't asking whether he could cheat. He was asking how to use a tool he already used, for purposes he genuinely didn't know were off-limits. And I, a teacher who had by then spent months using the exact same tool to draft rubrics and parent emails, had zero prepared answer. I was teaching with AI every week and I had never once taught about it.

Here's what I've come to believe: the moment you opened this book, you became a de facto AI literacy instructor. Not because anyone assigned you that role. Because your students are already using these tools — 59 percent of U.S. teenagers reported using AI tools for schoolwork in a 2024 Common Sense Media survey — and someone in the building needs to help them think critically about what those tools actually do. That someone is almost certainly you, whether your subject is English, biology, band, or pre-algebra.

This chapter is not about adding a unit to an already-overfull curriculum. It's about recognizing the AI literacy moments already embedded in your content and knowing how to use them.

You Don't Need a CS Degree. You Need a Good Question.

The most effective AI lesson I have ever run took eleven minutes. No laptops required. I read aloud an AI-generated paragraph about the causes of World War I — one I had prompted myself the night before — and asked my 8th graders a single question: "What is this paragraph leaving out?"

Within four minutes, a student named Priya had pointed out that the paragraph attributed the war almost entirely to the assassination of Franz Ferdinand and said nothing about imperial competition in Africa. Another student noticed the writing had no point of view, no argument — it just listed things. A third said it "sounds like a textbook but a boring textbook." That led to a conversation about what we expect writing to do that a summary can't. That conversation was better than any media literacy worksheet I had ever assigned.

The secret is this: AI outputs are perfect teaching texts because they are simultaneously impressive and flawed, and students can see both things at once. They're not defensive about criticizing the tool the way they sometimes are about criticizing each other's work. They'll dig in. They'll compete to find the error. And the skills they're building — sourcing, bias recognition, evaluating what's missing — are the exact skills that

define genuine intellectual work.

You already teach those skills. You just haven't used AI outputs as the practice text before.

Subject-by-Subject Entry Points

I want to be honest: I'm an ELA teacher. Everything below about ELA comes from my own classroom. Everything about other subjects comes from colleagues in my building who agreed to be pested with questions while I was writing this. I've run their lessons past them for accuracy. Still — you know your content better than I do. Take what works.

English Language Arts

ELA is where AI literacy fits most naturally, because the core question of ELA — what does this text do, and how does it do it? — applies directly to AI-generated writing.

Start with a comparison exercise. Generate a paragraph on any topic you're teaching — symbolism in The Giver, the structure of a counterargument, whatever — and put it next to a strong student paragraph from a previous year (anonymized). Ask: which one risks something? Which one has a voice? Which one would you want to read again?

Students almost always choose the human paragraph, and they can usually say why. That conversation is your opening. The follow-up question is the one that matters: "So what does writing need to do that this AI paragraph doesn't?"

You can also use AI to teach revision. Have students prompt an AI tool to write an introduction on a topic they know well, then revise the AI's draft into something better. The constraint forces them to articulate exactly what was wrong and exactly what they want instead. That's harder than revising their own drafts, where they're often too close to see the problems.

Social Studies

My colleague Daniel teaches 7th-grade U.S. history, and he's been running what he calls "the bias hunt" for about a year. He asks ChatGPT to summarize a historical event from a specific perspective — westward expansion from the view of a settler, for example — then asks it for the same event from a different vantage point. He prints both, puts them side by side, and asks students to identify what changed, what stayed the same, and what neither version mentions at all.

The exercise works because AI does something predictable: it shifts surface framing (word choice, who gets named, who's described as acting versus being acted upon) while often failing to fundamentally restructure the argument. Students can see the bias operating at the level of language rather than having to take Daniel's word for it.

He extends this into primary source analysis. Students compare an AI summary of a document to the document itself, looking for what the AI flattened, simplified, or got factually wrong. In one recent class, the AI's summary of a Reconstruction-era speech changed the speaker's actual claim. Three students caught it. That moment — that "wait, that's not what it says" moment — is worth more than a worksheet.

Science

The gift AI gives science teachers is confident wrongness. Ask any AI tool a sufficiently specific scientific question and it will, with high frequency, produce an answer that sounds authoritative and contains at least one error. This is not a bug for teaching purposes. It is the feature.

Keisha, who teaches 8th-grade earth science in our building, starts her unit on climate data by having students ask an AI tool three questions about climate projections. They're required to fact-check each answer against a named, verifiable source — usually NOAA or NASA's publicly available data pages. Almost every class finds at least one numerical error and at least one claim the AI presented as settled that is actually still contested in the research literature.

The lesson Keisha's teaching isn't "AI is bad." It's the same lesson she'd teach with any source: data claims require citations, and citations require verification. AI just makes the stakes feel higher because the tool sounds so sure of itself. Students find that tension interesting. They want to catch it.

Science also offers a natural hook for discussing how AI models are trained — which is a content-area question, not just a technology question. Where did the training data come from? Who produced it? What's left out? Those questions map directly onto the scientific method's concern with sample selection and data provenance.

Math

I'll be honest — math was the subject where I struggled most to see the entry point. Then Marcus's math teacher, a quiet and meticulous person named Trevor, showed me what he does, and I felt slightly foolish for not seeing it sooner.

Trevor asks students to use an AI tool to solve a word problem, then explain the AI's solution step by step — in their own words, out loud, to a partner. If they can't explain it, they don't understand it, and they can't use it. This isn't an AI lesson per se. It's a comprehension check with a new kind of source material. But it does something else: it shows students that using AI to get an answer is only the beginning of the work, not the end of it.

He also uses AI-generated "worked examples" with deliberate errors embedded — sometimes by AI's own mistake, sometimes ones he introduces himself — and turns error analysis into a weekly practice. Finding the flaw requires understanding the underlying mathematical reasoning. You can't fake your way through it.

The Arts

This one surprises people, but visual art and music teachers have some of the most immediate AI literacy opportunities in any building, because AI image generators and music tools make the question of authorship visceral in a way that text doesn't always.

My friend who teaches visual art at the high school level shows students AI-generated images alongside student work and asks two questions: Who made this, and what did they decide? The difference between a human artwork and an AI image often comes down to decision — the artist chose this angle, this palette, this crop, for reasons connected to lived experience and intent. AI doesn't decide. It predicts. Students in art class grasp this distinction quickly and with real feeling.

For music, the entry point is similar. Tools like Suno can generate a song in a specified genre in about thirty seconds. Play one for your class. Then ask: what's missing? What would a real composer have chosen that this didn't? The answers go deep, fast.

The SBLC Protocol

Across all of these subjects, I've found one analytical frame useful enough to write on my whiteboard and leave up for most of the year. I call it SBLC: Source, Bias, Limitation, Consequence.

It works like this. Any time students encounter an AI output — in class, assigned, or just in conversation — they can run it through four questions:

Source: Where did this information come from? Can I verify it independently? AI tools are trained on data scraped from the web as of a certain cutoff date. As of early 2026, ChatGPT's training data has a cutoff of early 2024; Claude's is similar. That means the tool may not know about events, studies, or developments from the past year or two. Students need to know this not to dismiss AI but to know when to double-check.

Bias: Whose perspective is centered here? Whose is missing? AI models reflect the biases present in their training data, which skews heavily toward English-language sources, Western viewpoints, and content produced by people with internet access. That's not a conspiracy. It's a data distribution problem. Students who understand this can ask better follow-up questions.

Limitation: What can't this tool do? What kind of question would break it? Numerical accuracy, recent events, local knowledge, cultural nuance — these are consistent weak spots. Naming the limitation specifically is more useful than a general "AI makes mistakes" warning, because it tells students where to focus their skepticism.

Consequence: What happens if someone acts on this output without checking it? This question moves AI literacy out of the abstract. A doctor who takes an AI-generated drug interaction summary at face value. A journalist who publishes an AI-summarized court filing without reading the original. A student who submits a bibliography the AI invented. Consequences make the stakes real.

You don't have to use all four questions every time. In a short class discussion, one or two is plenty. The goal is to make the frame habitual, so students reach for it automatically — not just with AI, but with any source

that presents itself as authoritative.

Run SBLC as a class discussion, a journaling prompt, a partner talk, a short written response. It transfers across grades and disciplines without modification. I've seen it used in a 5th-grade science class and a 12th-grade AP English class in the same week. Both worked.

Running a Class Analysis of AI Output

Here's a structure I've used and seen used successfully across subjects. Adapt the timing to your block length.

Step 1 — Generate together (5 minutes). Project an AI tool on the board. With the class watching, type a prompt related to your current content. Read the output aloud together.

Step 2 — First reactions (3 minutes). Ask: "What's right? What seems off? What surprised you?" Take hands or do a quick turn-and-talk. Don't evaluate their responses yet — you're just surfacing observations.

Step 3 — Run SBLC (10–15 minutes). Work through one or two of the four questions as a full-class discussion. Assign specific claims in the AI output to small groups to verify or challenge using their textbook, a named database, or a primary source.

Step 4 — Synthesize (5 minutes). Ask the class to agree on one sentence that describes what this AI output did well and one that describes what it got wrong or left out. Write both sentences on the board.

Step 5 — Connect to content (2 minutes). Close by naming the skill explicitly: "What you just did — checking a source, finding what's missing, asking who's centered — that's exactly what historians do. What scientists do. What readers do."

Eleven minutes to thirty-five minutes, depending on how deep you go. No laptops required unless you want them for the verification step. The conversation itself is the lesson.

Ready-to-Run Discussion Questions

These work for any subject, any grade. Use one at the start of a lesson, or as an exit ticket.

- "What would you need to know about a source before you trusted it? Does that change with AI?"
- "If an AI got this answer right, does that mean it understands the concept? How would you tell?"
- "What question would this AI be most likely to get wrong? Why?"
- "Who is missing from this AI's answer?"
- "If you used this output without checking it, what's the worst thing that could happen?"

Student AI Literacy Rubric

This rubric is designed for a short written or oral response where students analyze an AI-generated text. Four dimensions, four levels. Adapt the language for your grade band.

| Dimension | Beginning | Developing | Proficient | Advanced | |---|---|---|---|---| | Source awareness | Does not identify that AI has a training cutoff or data source | Names that AI has limitations but can't describe them specifically | Identifies the training cutoff and explains how it affects this specific output | Identifies limitations, finds at least one specific claim that may be affected, and proposes a verification strategy | | Bias recognition | Does not identify whose perspective is centered or absent | Notes that the output has "a perspective" without specifics | Names whose viewpoint dominates and whose is missing, with at least one textual example | Explains why that bias likely exists (data distribution, language dominance) and what a more complete version would include | | Limitation analysis | Accepts AI output at face value or rejects it entirely without reasoning | Names a general limitation ("AI makes mistakes") | Identifies a specific type of limitation (numerical, temporal, cultural) relevant to this output | Provides evidence of the limitation by comparing the AI output to a verifiable source | | Consequence thinking | Does not consider what happens if the output is used unchecked | Identifies that using wrong information "could cause problems" | Describes a specific realistic consequence relevant to the content or context | Proposes a concrete verification or correction strategy and explains when the output is safe to use as-is |

Here's what I want to leave you with. You are not being asked to become a technology teacher. You are being asked to do what you've always done — help students think carefully about texts and claims and sources — with a new kind of text in the room.

AI outputs are not neutral. They reflect choices: what data was included in training, what languages dominated, what questions the tool was optimized to answer. Teaching students to see those choices is the same work as teaching them to see whose voice is centered in a novel, whose data gets used in a scientific study, whose story gets told in a history textbook.

Marcus, by the way, ended up writing one of the strongest argument essays in that class. He did use AI — to brainstorm counterarguments, which we had agreed was fair game. His final draft had a voice that was unmistakably his. Impatient, funny, precise. No AI would have written it.

That's the point. That's what you're protecting when you teach this.

CHAPTER XII

Your First 30-Day Teacher AI Plan

The first time I sat down to deliberately practice using AI — not desperately, not at midnight with a pile of ungraded essays, but intentionally, with a cup of coffee and thirty free minutes on a Tuesday afternoon — I felt ridiculous. Like I was learning to use a microwave at age thirty-four. What was there to learn? You open it, you press a button, things heat up.

That is not how it works. That is absolutely not how it works.

What I discovered in those thirty minutes, and what I kept discovering over the following month, is that AI is closer to a new instrument than a new appliance. A microwave works the same whether you've used one before or not. A guitar does not. The first week you pick one up, you're mostly making noise. By week four, if you've been deliberate about it, you're playing something recognizable. By week twelve, muscle memory starts to take over and you stop thinking quite so hard about where your fingers go.

This chapter is your first four weeks. Not your whole journey — just the part where you learn where to put your fingers.

I'm going to ask one thing of you before we start: pick a specific week on the calendar right now, not "next week sometime," and commit the thirty-day window. Block four Friday afternoons, even five minutes each, in your planner or your phone. Because the teachers I know who've made AI stick didn't stumble into a habit. They scheduled one.

Before Week 1: The Only Setup That Actually Matters

You do not need to configure seventeen settings or watch a two-hour tutorial. You need to do one thing: write a system prompt describing you and your classroom, and save it somewhere you can paste it in under ten seconds.

Mine lives in a note called "AI context" on my phone. It's three sentences:

"I am an 8th-grade ELA teacher at a public school in Columbus, Ohio. My students are a mix of proficiency levels; about 30% have IEPs or 504 plans, and 20% are English language learners. I follow Ohio's Learning Standards and I want all output formatted for direct classroom use without corporate jargon."

That's it. Those three sentences save me from having to re-explain myself every single session. They turn generic AI output into something that has at least a fighting chance of being useful in my actual room with my actual kids.

Paste your version into whatever tool you're starting with — Claude at claude.ai, ChatGPT at chatgpt.com, whichever — at the beginning of every new conversation this month. After thirty days, check whether the habit is automatic. It usually is.

Week 1: One Task. Just One.

The biggest mistake new AI users make — and I made it, hard, in October of my sixth year — is trying to rebuild their entire workflow at once. They come in hot, full of ambition, and they ask AI to rewrite their unit plans and overhaul their rubrics and draft twenty parent emails before Friday. Then the output is mediocre on three of those things and they conclude AI doesn't work and go back to their Sunday-night spreadsheet.

Start smaller than you think you should.

Week 1 has one job: pick a single low-stakes task you already do regularly, and do it with AI instead of doing it alone. Two options work especially well here because the stakes are genuinely low and the time savings are immediately visible.

Option A: The Parent Newsletter

Most teachers send some version of a weekly or biweekly "what's happening in Room ___" email. It's not technically hard. It's just one more thing that takes fifteen to twenty minutes you don't have. This is a perfect first prompt.

Try this one, verbatim, after pasting your context note:

> *"I need a short parent newsletter for [your subject], [your grade], covering the week of [dates]. Here's what we did: [paste your three or four bullet-point notes from the week]. Make it warm, about 150 words, and end with one specific thing parents can ask their kid about at dinner."*

The output will not be perfect. You'll change a word here, cut a sentence there. That's fine. Time yourself. Most teachers report spending four to seven minutes on the revision rather than fifteen to twenty on writing from scratch. That's your proof of concept, and that's all you need this week.

Option B: One Quiz

Pull a learning objective from next week's plans — one specific, measurable one — and try this:

> *"Generate a 10-question multiple-choice quiz on [topic] for [grade level] students. Include two questions at the knowledge level, four at the comprehension level, and four at the application level. Write four answer choices per question and mark the correct answer. Flag any question where the distractors might be ambiguous."*

Read every question before you use it. That's not optional. But you'll almost certainly find seven or eight are solid, one or two need a rewrite, and you've built something in twelve minutes that would have taken forty-five.

One Thing to Notice This Week

Where did the AI surprise you — in a good way or a bad way? Don't evaluate the whole technology. Just notice one specific moment when the output was better than you expected, or one specific moment when it missed something obvious. Write it down. That observation is data you'll use in Week 4.

Friday Reflection Question (5 minutes)

What would this week have cost me in time if I hadn't tried AI on that one task?

Be honest. Not aspirational, not self-deprecating. Just calculate the actual minutes. This number matters.

Week 2: Lesson Planning Gets Its Turn

By the start of Week 2, you've done AI once in a real context and lived to tell the tale. Now you're going to do something scarier: bring it into lesson planning, which is the task that probably consumes more of your Sunday than anything else.

The goal this week is to use AI to produce a complete five-day lesson arc — not to copy it wholesale, but to see what a starting draft looks like and how long your "make it mine" editing pass actually takes.

Here's the prompt that gets me from a blank page to a usable skeleton in under twenty minutes:

> *"I'm planning a five-day unit for [grade] [subject] on [topic]. The key standard I'm addressing is [paste the standard text]. Students will produce [final product or assessment] at the end of the week. For each day, give me: a learning objective, the main activity, a discussion question, and a two-question exit ticket. Format it as a table."*

What comes back will be recognizable but generic. That's expected. Your editing pass — which I described in Chapter 3 as the "make it mine" pass — typically runs eight to twelve minutes and involves three moves: replacing the generic exemplar with one from your actual classroom, swapping the generic discussion question for the real argument you want students to actually have, and cutting Day 3 or 4 to fit the reality that something always goes sideways and you'll need the buffer.

If you have a student teacher or a planning partner, do this together. Watching someone else prompt AI for the first time is instructive in ways that are hard to describe — you see the instinct to over-explain, or to ask for too much at once, or to accept the first output without pushing back. Prompting is a skill that sharpens fastest when observed.

One Thing to Notice This Week

How much of the AI draft did you actually keep? Track it roughly. Some teachers keep 40%, some keep 80%. Neither number is right or wrong — but knowing your own ratio tells you something about how much scaffolding you need the AI to provide versus how much you're building yourself.

Friday Reflection Question (5 minutes)

Did this week's planning feel different? If yes, how? If no, what got in the way?

This is not about whether AI "worked." It's about your own relationship to the planning process. Some teachers feel immediate relief. Others feel vaguely guilty, like they're outsourcing something that should be harder. Both reactions are real and worth examining.

Week 3: Go Deeper on Differentiation or Documentation

You've got two options for Week 3, and the right one depends on what your job is actually made of. If differentiation is your heaviest lift — if you spend hours producing tiered materials — go that direction. If IEP and 504 documentation is what's quietly bankrupting your October, go there instead.

Path A: Differentiation

Take a single text or task you're already planning to use next week. Paste it in — or describe it in specific terms — and try this:

> *"Here is a [genre/type] passage I'm using with my [grade] class. Please produce three versions: one at approximately the 6th-grade Lexile level, one at grade level (8th grade), and one extended version that adds a layer of analytical complexity for students reading above grade level. Keep the core content and argument identical across all three versions. For the lowest-level version, also add five vocabulary scaffolds in bold."*

Then add, in a follow-up message: "Now generate a sentence frame bank of eight frames that struggling writers could use to respond to this text."

Two prompts. Under ten minutes. You will need to read the Lexile-adjusted versions carefully — AI's sense of what constitutes "simpler" language is imprecise and sometimes condescending. Fix the parts that flatten the content too aggressively. But you will not be starting from a blank page, and that is the entire point.

Path B: Documentation

Pull out a rough observational note from the past two weeks — the kind of thing you typed into your phone between passing periods, half-formed, shorthand. Something like: "Marcus — still not initiating during partner work, reads at grade level per last fall's assessment, struggles to organize ideas in writing, strong verbally."

Try this:

> *"I am writing a present-level-of-performance summary for a student's IEP. Here are my rough observational notes: [paste notes]. Please transform this into two professional paragraphs suitable for an IEP document — one describing academic performance and one describing functional performance. Use objective, observable language. Do not diagnose or speculate beyond the notes*

I've provided."

Review every word before it goes anywhere. Check it against your district's preferred language. Flag anything that feels like extrapolation beyond what you actually observed. This is a draft, and you are the professional who signs off. But getting from rough notes to structured language in four minutes instead of twenty-five is real. I've done it 200 times now and I still feel the relief.

One Thing to Notice This Week

Where did AI save you the most time — and where did it create new work? Because sometimes it does create new work. A poorly calibrated Lexile adjustment means you're editing more than you would have written. A documentation draft that uses words outside your district's style guide means extra revision. Notice the friction points. They're telling you something about how to refine your prompts next time.

Friday Reflection Question (5 minutes)

Which part of this week's AI-assisted task required the most human judgment?

This is the question that keeps AI honest. The answer is always something — and identifying it is how you figure out where your expertise is irreplaceable.

Week 4: Lock In What's Working, Let Go of What Isn't

Week 4 is not a week for learning new prompts. It's a week for auditing the previous three.

Look back at your Friday reflections. Look at the time you calculated at the end of Week 1. Think about Week 2's question — did the planning feel different? — and Week 3's question about where human judgment was irreplaceable. Now answer three questions in writing:

- Which one or two AI workflows genuinely saved me time in a way I could feel by Friday?
- Which workflow felt forced, or produced output I didn't really trust?
- If I kept only two AI habits permanently, what would they be?

Most teachers land on two things. Two, not eight. The parent communication workflow. The quiz draft. The IEP progress note. Lesson planning on Sunday afternoon. Differentiated exit tickets. Whatever yours are, name them, and make them official: those are your AI habits now. Everything else is optional.

This week, run your two chosen workflows one more time each — not to learn them, but to do them fast. Speed is the goal. The first few times you built the house; now you're learning how to live in it without thinking about the floorplan.

One Thing to Notice This Week

How long did each workflow take compared to Week 1? It will almost certainly be faster. That's the instrument analogy again — repetition builds muscle memory, and muscle memory is what turns a promising

skill into an actual time saver.

Friday Reflection Question (5 minutes)

What will I protect with the time I'm reclaiming?

Not "what will I do with the extra time at school." What will you protect in your life outside school. This question is the most important one in the book, and I mean that.

The Real Numbers

My first year of teaching, I tracked my work hours for three weeks out of morbid curiosity. The average was 14.2 hours of planning and prep per week, not counting actual instruction time. I was twenty-four years old and I thought that was just what teaching was.

By my fifth year — before AI — I had gotten it down to about eleven hours through sheer accumulated efficiency: I knew the curriculum cold, I'd stopped reinventing every unit, I had a folder of go-to materials. Eleven hours is still too many if it's coming out of Sunday, but it felt like progress.

Now I average nine. Some weeks eight. A brutal IEP documentation week might push me back to eleven.

The difference between eleven and nine — two hours a week — is 72 hours a school year. That's almost two full work weeks. I want you to actually sit with that number, because the instinct is to think "only two hours?" and feel underwhelmed. But 72 hours is the ten-day trip I took to Portugal in July. It's the Saturday mornings I stopped spending at the kitchen table with a laptop. It's showing up to class on Monday with something left in the tank.

AI didn't give me my life back. That's an overclaim and I won't make it. What it did was plug a specific category of leak — the administrative drafting work, the first-pass formatting, the 11 p.m. panic about not having an exit ticket — so that my actual teaching energy could go toward the parts of this job that are genuinely mine and genuinely irreplaceable.

Nobody has built an AI that can look at a ninth-grader in the third row who hasn't spoken in two weeks and know what she needs. Nobody has built one that can hear a student read aloud and understand precisely where his fluency broke down and why. Nobody has built one that can hold the room on a Friday in March when everyone is cooked and you pivot to something that somehow works. That's you. That has always been you.

The risk of this moment — and I say this as someone who thinks these tools are genuinely useful — is that the conversation about AI in schools consumes energy that should be going toward the harder, slower, less legible work of teaching actual human beings. Every hour a teacher spends anxious about AI detection, or chasing the newest model release, or rebuilding a workflow that was already fine, is an hour not spent on a kid.

So here is the final instruction, and it is the simplest one in this book:

Decide, in advance, which hours belong to your job.

Not loosely. Not aspirationally. Actually decide. Write down the number. Mine is eight hours of planning and prep per week. When the week hits eight hours, I stop. Sometimes a deadline means I go over. But I know exactly when I've crossed the line, and I know it's a choice, not an accident.

AI is one of the tools that makes that boundary possible to hold. The boundary itself is the point.

Your thirty days start when you're ready. Go small, go slow, notice what's true, and protect the hours that were never the school's to begin with.