

AP Research: Using AI Ethically in Source Diving

A student guide to finding and using sources ethically in the age of AI

The Official College Board Policy (AP Capstone, verbatim)

The heart of the policy is what AI may and may not do. That is the paragraph in the highlighted box below. Read that part most closely.

Generative AI tools must be used ethically, responsibly, and intentionally to support student learning, not to bypass it. Accordingly, all performance tasks submitted in AP Seminar and AP Research must be the student's own work. While students are permitted to use generative AI tools consistent with this policy, their use is optional and not mandatory.

THE CORE OF THE POLICY

Students can use generative AI tools as optional aids for **exploration of potential topics of inquiry, initial searches for sources of information, confirming their understanding of a complex text, or checking their writing for grammar and tone.** However, students must **read primary and secondary sources directly, perform their own analysis and synthesis of evidence, and make their own choices on how to communicate** effectively both in their writing and presentations. **It remains the student's responsibility to engage deeply with credible, valid sources and integrate diverse perspectives** when working on the performance tasks. Students must complete interim "checkpoints" with their teacher to demonstrate genuine engagement with the tasks.

The rest of the policy, in brief: students must complete interim checkpoints to show the work is genuinely theirs (in AP Research these are in-progress meetings and the Process and Reflection Portfolio, PREP), and failure to complete them can result in a score of 0. Teachers affirm that the work is authentic, and the College Board may investigate and request a student's interim work where misuse is suspected.

Source: College Board, AP Central, "Guidance for Artificial Intelligence Tools and Other Services," AP Capstone Policy, 2025–26 (apcentral.collegeboard.org). The highlighted paragraph is quoted verbatim; the brief summary above paraphrases the policy's checkpoint and affirmation provisions.

The One Rule That Matters

The College Board lets you use AI for **finding and understanding** sources.

It does **not** let AI **read, analyze, or write** for you.

If you submit work AI created, it can lead to a **score of zero** on that task. Final submissions run through Turnitin for AI and plagiarism detection, and the College Board can investigate a submission and request your interim work to confirm the engagement was genuinely yours. Skipping the required checkpoints can itself result in a zero on that task.

Did the AI help me find or understand a source?

Or did it do my thinking and writing for me?

The first is allowed. The second is a violation.

Finding Your Sources

The goal is to find **academic sources** related to your working question.

AI and the databases below are not competing methods, they are two ends of the same pipeline: **AI points, databases confirm.**

AI is fast but can be unreliable. It gives you leads: researcher names, search directions, sometimes a real link and sometimes an invented one. A database is slower but authoritative. It is where you confirm a source is real, get the correct citation, and actually open and read the paper. You will bounce between the two: ask AI for a direction ("who are the key researchers on this?"), then go into a database to find their actual work, verify it, and read it.

That is why you still need to know which databases to use even when AI is your starting point.

Where to confirm and read (click to open):

- **Google Scholar** (free, broad, best starting point) scholar.google.com
- **ERIC** (free, education-focused, ideal for school and learning topics) eric.ed.gov
- **PubMed** (free, anything health, medical, or psychology related) pubmed.ncbi.nlm.nih.gov
- **Semantic Scholar** (free, AI-assisted discovery) semanticscholar.org
- **CORE** (free, open access papers) core.ac.uk
- **DOAJ (Directory of Open Access Journals)** (free, full-text articles you can always open) doaj.org

- **JSTOR, EBSCO, or ProQuest** (strong for humanities and social science; access through a library subscription) [jstor.org](https://www.jstor.org)

GREEN: allowed, use freely

- Brainstorming or narrowing a **topic** of inquiry
- Running an **initial search** for where to look, including asking a web-connected AI for links to real sources
- Asking AI to **explain a complex text (for example, asking AI to explain a dense abstract)**, so you know whether to pull the paper
- Generating **search terms and keywords** to use in your databases
- Checking your own writing for **grammar and tone** (checking, not rewriting)

RED: not allowed, will cost you the score

- Letting AI **write your annotations** or summarize a paper you didn't read
- Submitting an **AI-generated literature review** or any AI-written text as your own
- Citing a source you found through AI but **never actually opened**
- Having AI **analyze or synthesize** the evidence for you
- Using AI to **rewrite** your sentences (vs. flag a grammar issue)

“Can I just ask AI for the link?”

Yes, you can. Asking a web-connected AI for a link falls under what the College Board calls **“initial searches for sources of information,”** as long as you then verify and read the source yourself. A link from AI is a **lead, not a citation**. Two things are true no matter where the link came from: you have to open it and confirm it is real, and you have to read it yourself before it goes in your bibliography. The trap is that an AI without live web access will invent a link that looks perfect and goes nowhere. So the habit is simple: **ask for links if you want, but treat every link as guilty until you open it and verify it.**

Sharpen your prompts: add constraints

A vague prompt gives you vague, unusable results. Add constraints so the leads you get back are ones you can actually use for a class like this. Two that matter most:

- **“The link must be accessible.”** Ask for sources that are free to read or open access, so you are not stuck behind a paywall. A source you cannot open is a source you cannot read, and one you cannot read cannot go on your list.
- **“Keep it reasonable in length.”** Ask for journal articles in the range of a normal paper, not 200-page books or dissertations. You don't need a 200 page book, especially when your looking at a 10-20 source to use.

Example of a sharpened prompt: *“Suggest 4 to 5 peer-reviewed articles on teen sleep and academic performance. Prefer sources that are free to access, and keep them to normal article length rather than full books. Give me real citations and links so I can verify myself.”*

Steer AI toward credible, peer-reviewed sources

AI will hand you blogs, news articles, and student essays right alongside real scholarship unless you tell it not to. You cannot fully trust it to filter for quality, but you can push it hard in the right direction with a few moves:

- **Say “peer-reviewed” explicitly.** Ask for “peer-reviewed journal articles,” not just “sources.” It is the single most useful word you can add.
- **Name the kind of study.** Ask for “empirical studies, systematic reviews, or meta-analyses.” These terms point AI at real research rather than opinion pieces.
- **Point it at trusted homes.** Ask it to prefer sources indexed in Google Scholar, PubMed, ERIC, or JSTOR, or published in academic journals. You can even name a database: “only suggest things I could find on PubMed.” This also depends on what type of subscriptions your school has (or you have individually paid for).
- **Ask for the evidence of credibility.** Ask AI to tell you the journal name and whether it is peer-reviewed for each source. Then you confirm it yourself, because AI can be wrong about this too.
- **Exclude what you don’t want.** Add “no blogs, news sites, or summaries” to cut the noise

Stronger prompt with credibility built in: *“Suggest 3 to 5 peer-reviewed journal articles, empirical studies or meta-analyses, on teen sleep and academic performance. Prefer things indexed in PubMed or Google Scholar, free to access, and normal article length. For each one, give me the journal name, the citation, and a link so I can verify it myself. No blogs or news articles.”*

One honest caveat: even with the perfect prompt, AI is not a reliable judge of whether something is truly peer-reviewed. These moves raise the quality of what comes back, but the final check is always yours: open the source, find the journal, and confirm it is the real thing.

Prompt Bank: How to Actually Ask AI

Each prompt uses one topic, “**Does sleep duration affect academic performance in teenagers?**” and explains *why* it is allowed or not. Understanding the *why* is what lets you make the right call on a prompt nobody handed you. The pattern: **AI, then you read, then AI, then you write.**

Step 1: Sharpen your search

“I’m researching whether sleep duration affects academic performance in teenagers. What are the key academic fields, search terms, and synonyms I should use to find peer-reviewed sources on Google Scholar and similar databases? List the terms so I can search them myself.”

What you might get back: fields like adolescent psychology, chronobiology, education research; terms like sleep duration, academic achievement, adolescents, sleep quality, sleep deprivation.

WHY this is allowed: you are not asking AI for sources, you are asking it to translate your everyday question into the vocabulary scholars use. The policy permits AI for initial searches for sources of information. Generating search terms tunes the search; you still run it and you still find the sources.

Verify: catch the fake citations (use this every single time)

“Does this source actually exist? Give me the DOI, journal name, volume, and year so I can locate it myself.”

WHY this matters more than any other step: **AI tools can hallucinate.** They generate citations and links that look flawless but describe a paper that was never written, because they predict what a citation should look like, not whether it exists. If you cite a source you never opened, either it is fake and your reviewer finds out, or it is real but says something different from what you claimed. **Both are fatal in AP Research.** If you can’t pull it up yourself and open it, it does not exist for your purposes.

Step 2: Decide if a paper is worth your time

“Here is an abstract I found in a database: [paste the abstract]. In plain English, what is this

study claiming, what method did it use, and is it actually about teen sleep and grades? I'm deciding whether to read the full paper."

What you might get back: "This study surveyed 5,000 high schoolers, measured self-reported sleep against GPA, and found a weak positive association. It is about your topic but relies on self-reported data." This can be useful if the abstract is very long.

WHY this is allowed: this is the policy's confirming their understanding of a complex text. Abstracts are dense; using AI to decode one so you can decide whether a paper is worth the 20 to 30 minutes it takes to read is a comprehension aid. The boundary is that it tells you whether to read, it does not read for you. The output is a triage decision, not bibliography material.

Step 3: READ THE PAPER. No AI. (this step is the whole point)

There is no prompt here, and that is deliberate. Open the source. Read it. Take your own notes on what it argues, how, what its limits are, and how it connects to your question.

WHY there is no AI here: this is the line the entire policy is built around. The College Board requires you to read primary and secondary sources directly and perform your own analysis and synthesis of evidence. Reading is the thing AP Research exists to assess. Your annotation, your literature review, and your whole paper are evidence that you engaged with the scholarship. If AI read it instead of you, that evidence is fake, and you will expose it yourself the moment a checkpoint question goes one layer deeper than the AI summary did.

Step 4: Pressure-test your own understanding (after you've read)

"I read a study arguing that teens who sleep under 7 hours have lower GPAs. I think its main limitation is that sleep was self-reported, which might be inaccurate. Am I understanding the concept of self-report bias correctly?"

WHY this is allowed: notice who is doing the thinking. You read the study, you identified the limitation, you formed an interpretation. AI is only checking whether you understand a concept correctly, like asking a teacher "did I get this right?" The instant you flip it to "what are the limitations of this study?" and let AI generate the analysis, you have crossed into having AI perform analysis and synthesis, which is prohibited. **Same paper, same tool, opposite sides of the line. The difference is who generated the idea.**

Prompts That Cross the Line, and Exactly Why

Prompt	Why it is a violation
<i>"Summarize this paper for my annotated bibliography."</i>	The annotation IS the assignment. AI writing means AI creating your work. Automatic zero.
<i>"Write a literature review on sleep and academics."</i>	This is the performance task itself. Letting AI write it is the clearest possible violation.
<i>"Rewrite this paragraph to sound more academic."</i>	The policy allows checking grammar and tone. We read that as checking, not rewriting: rewriting hands your voice and choices to AI, which the policy reserves for you.
<i>"What are the limitations or themes across my sources?"</i>	Synthesis across sources is the core intellectual work being assessed. It has to be yours.

The pattern in all four: they each ask AI to produce a piece of the assignment, the words, the analysis, or the synthesis. The green prompts ask AI to help you find or understand so that you can then do the producing.

That single distinction, **help me think vs. think for me**, is the whole rule, and it is the test you apply when you hit a prompt this handout didn't predict.

Worked Example: Running the Process for Real

Here is the whole workflow on one real source, start to finish, so you can see what good practice actually looks like. We stay with the question

“Does sleep duration affect academic performance in teenagers?”

1. We asked AI for leads (allowed)

“I’m researching whether sleep duration affects academic performance in teenagers. Suggest a few peer-reviewed journal articles or meta-analyses, roughly 10 to 25 pages, and give me the journal name and a real citation and link for each so I can verify them myself.”

AI suggested, among others, a 2010 meta-analysis by Dewald and colleagues on sleep and school performance. That is a lead. On its own it is worth nothing until we verify it.

2. We verified it was real (the critical step)

We searched the citation to confirm it exists and that the details are right. It checks out:

Verified record:

Dewald, J. F., Meijer, A. M., Oort, F. J., Kerkhof, G. A., & Bögels, S. M. (2010). The influence of sleep quality, sleep duration and sleepiness on school performance in children and adolescents: A meta-analytic review. *Sleep Medicine Reviews*, 14(3), 179–189. DOI: 10.1016/j.smrv.2009.10.004. PubMed ID: 20093054.

Find it yourself: pubmed.ncbi.nlm.nih.gov/20093054/

Because the paper is real and findable, it earns a place on the list. If the search had turned up nothing, we would have dropped it, no matter how confident the AI sounded. (The official journal version is paywalled, but a free copy is often available through Google Scholar or your school or library databases, so search the title before assuming you cannot read it.)

3. We read it ourselves, and caught something

Reading the actual findings is where the real work happened, and it changed the story. Here is the paper’s abstract to get you oriented before you read the full text:

Abstract (Dewald et al., 2010):

Insufficient sleep, poor sleep quality and sleepiness are common problems in children and adolescents being related to learning, memory and school performance. The associations between sleep quality (k=16 studies, N=13,631), sleep duration (k=17 studies, N=15,199), sleepiness (k=17, N=19,530) and school performance were examined in three separate meta-analyses including influential factors (e.g., gender, age, parameter assessment) as moderators. All three sleep variables were significantly but modestly related to school performance. Sleepiness showed the strongest relation to school performance ($r=-0.133$), followed by sleep quality ($r=0.096$) and sleep duration ($r=0.069$). Effect sizes were larger for studies including younger participants which can be explained by dramatic prefrontal cortex changes during (early) adolescence. Concerning the relationship between sleep duration and school performance age effects were even larger in studies that included more boys than in studies that included more girls, demonstrating the importance of differential pubertal development of boys and girls. Longitudinal and

experimental studies are recommended in order to gain more insight into the different relationships and to develop programs that can improve school performance by changing individuals' sleep patterns.

A lazy researcher would assume a sleep study concludes “more sleep means better grades.” Read the abstract closely and the real story is more interesting:

- All three sleep measures related to school performance, but only modestly.
- Sleepiness had the strongest link ($r = -0.133$), then sleep quality ($r = 0.096$).
- Sleep duration was the weakest of the three ($r = 0.069$), which is the opposite of what most people assume.

Why reading mattered here: the source you started with to prove “duration drives grades” actually shows duration is the weakest predictor of the three. You only learn that by reading it. An AI summary pasted into your bibliography would have sent your whole research question in a shaky direction, and a checkpoint question would have exposed that you never read the paper.

4. We checked our own understanding (allowed)

“The Dewald meta-analysis reports sleep duration as the weakest predictor ($r = 0.069$) and sleepiness as the strongest ($r = -0.133$). I think this means how tired a student feels matters more than raw hours in bed. Am I reading those correlations correctly?”

We brought our own reading and interpretation, and used AI only to confirm we understood the statistics. We did not ask it what the paper means or to write our annotation.

5. We wrote the annotation ourselves

Student-written annotation (in your own words):

Dewald ran three separate meta-analyses (each covering well over 10,000 students) and found that sleep relates to school performance only modestly. The surprising part for my question is that sleep duration is the weakest of the three measures, behind sleepiness and sleep quality. This pushes me to rethink my working question: maybe I should be measuring how rested students feel, not just how many hours they sleep. It also makes me question about self-report may inflate these links, which I should watch for in my own survey.

That annotation could only be written by someone who read the paper and connected it to their own project. That connection is exactly what AP Research is grading, and exactly what no AI can do for you.