

Annotation and the power of teaching a process

Karl Seeley
Hartwick College *

August 13, 2026

Abstract

This paper reports on an informal experiment in an economic history course in the Spring 2026 semester. I used physical annotation on a printed text as the key component of a structure for simultaneously avoiding over-reliance on LLM's and fostering stronger student engagement with course material and deeper learning. Compared to other pedagogical approaches, this system gets the teacher closer to directly observing the process of ultimate concern: the allocation of attention and the inner response to ideas. The approach is consistent with findings from cognitive science on the relationship between mental effort and learning, on the importance of fact-based knowledge for thinking, and on the role of attention as a scarce resource that people must learn to apply and allocate. The initial results warrant further exploration.

Having never been a huge reader in general, let alone one who takes the time to annotate, this class has progressed me in both of those aspects. I hope to be able to continue utilizing these habits into my future so I can be the most successful self I can be.

Student
ECON 311

1 Introduction

The eruption of large language models (LLMs) into the public space has created a raft of challenges for educators at various levels, including post-secondary. We face increasing difficulty in discerning when class work reflects a student's effort and engagement, as opposed to having let the machine do it. This obviously raises new difficulties in evaluating student work, but many educators are also

*I am grateful to Kristin Jones and Carli Ficano for useful feedback on an earlier version.

grappling with a deeper issue of how we should be structuring class time and assignments so as to foster real learning in our current environment.

This paper presents one particular response to those challenges, based around physical annotation of a text and discussion with students, either individually or in small groups of two to four students; a second prominent component is the composition of reflection paragraphs after small-group discussions of readings, or following a lecture-based classroom period. This approach moves a lot of focus off of the student's finished work and onto the process that comes to fruition in that work.

Arguably, this would have been a constructive change at any earlier point in my 24 years as a professor, as it has led to class discussions and in-class written answers that show more engagement and competence than I have seen in many years. But LLMs forced my hand by making it radically easier for students to submit the "product" without having done any of the process.

I start by specifying my goals as a teacher, and then describe in detail the approach used in a recent course on economic history. After that, I present the outcomes and review some research on learning and cognition to explain how the partial exclusion of LLMs is important to achieving my teaching goals. I end with a further discussion of findings from cognitive science suggesting why the approach described here may be an effective tool.

2 Goals in teaching

It is easy to state my goals as a teacher in short form, as done below in Sec. 2.1. If you want a review of the cognitive theory and view of AI that have influenced me, make your way through Sec. 5.1 and 5.2 as well. If that material is already familiar or it bores you, skip to the summary statement of my objectives in the last paragraph before Sec. 3.

2.1 Goals

When you successfully complete a class of mine, I want you to understand something about the world that you didn't before; since I'm an economist, that usually means some aspect of economics. I also want you to retain a useful piece of what you have come to understand. If I were to ask a student 10 years out what they remembered from my course, I would expect it to be a small fraction of the detail they had had mental access to a decade earlier. At the same time, I want this alumnus/alumna to still have things in their mind that they first encountered in my course, or developed a greater understanding of during my course, and that are still useful when they look at the world and navigate life.

It's common to talk about education as developing students' ability to think. I share that goal, but over the decades of my time as a student and then as a professor (and thus, in some sense, perpetually a student), I have found a

deepening appreciation for the role of fact-based knowledge as a substrate for the ability to think and understand. This implies that, as part of my goal of helping you understand something about the world, I also want you to simply know some stuff.

Section 5.1 addresses the importance of this on-board knowledge, while section 5.2 looks at how LLMs affect the acquisition of such knowledge. For now, the upshot is that to reach my fundamental goal of increasing students' lasting understanding of the world, I have an instrumental goal, in many contexts, of limiting their use of AI, if not boxing it out entirely—not merely in the classroom, but outside it as well.

The next section presents my current way of doing that.

3 The approach

3.1 Annotation

The core activity in my Spring 2026 class in Economic History was physical annotation on the physical page. For a scholarly article or piece of journalism, the student prints out the piece. For a book, the student buys the book.¹ The instructions to the students evolved over the term, but a late version was:

“Read with a pen or pencil in hand. The goal is not to note every important thing in the text, nor every connection that you see—there’s only so much you can take in on a first pass. But in a book, you usually shouldn’t go more than a couple of pages without noting something. In the format of a scholarly article (more words per page than most books, often denser content as well), you should probably have one or two notations per page.

“Notations are comments in the margins or other available space on the page, and they fall into two categories.

1. Content notations Simply noting a particular piece of content that seems important to have ready access to later: a fact, an argument, an idea. The notation itself can be anything from a word to a sentence; enough to be meaningful (in the context of the chapter or paper), but no longer than it needs to be.
2. Reflective notations Capturing your own thought in response to what you’re reading. Sometimes this is a question, e.g., Does this imply X? It can be a connection to something else you know (whether from elsewhere in the course, or from anywhere at all). It can be expanding on an idea in the text, or disagreement with an idea in the text (say *why* you disagree).

¹In this course, the book was *Slouching toward utopia* (DeLong, 2022), available as a trade paperback for about \$18, so not a massive expense.

“Annotation may include underlining or highlighting, to draw attention directly to the specific passage that prompted a notation. Or a vertical line may be used alongside a longer passage such as a paragraph, to indicate that a notation applies to that entire stretch of text. It is OK if *occasionally* an underline or vertical line stands alone, but for two reasons, the default should be for annotations to include your own words you have set on the page:

1. A major purpose of the work you’re doing is to turn the text into a tool for *your* use. Say there’s a paragraph about food shortages in the Netherlands during World War II. If you just put a vertical line next to the paragraph, you have to re-read some of the paragraph to both see what the paragraph is about and (hopefully) remember why you thought it was worth noting. If you write, “Hunger in Netherlands late in war” in the margin, you will know the relevance as soon as your eye lands on those words.
2. A second major point of this work is to create “hooks” to help the material in the text better stick in your mind—you’re creating a Velcro pad in your mind, and leaving little Velcro hooks all over the text so that the two will stick together. Crafting a phrase or a sentence is more work than simply underlining. There is effort required to collapse the paragraph down to, “Hunger in Netherlands late in war.” There is work in taking a feeling about the text and turning it into a specific sentence expressing your question, speculation, agreement, or objection. That work is what is making the Velcro hooks.

“When you’ve finished reading and annotating, look back over your work. Write down at least one overall response: a summary statement, a particularly important point from the reading, a particularly important question prompted by the reading. These overall responses should arise from your annotations—they should be the result of what you noted.”

In addition to written instructions, the first week of class included a mid-length, accessible reading (Wyman (2025)) for students to print out and annotate. That gave them practice at doing it and allowed me to give feedback and clarify expectations.

3.2 Discussion/assessment

The purpose of the annotation activity is to goad the student into substantive engagement with the text on their own. The assessment of this engagement involves scanning the reading as a whole, looking closely at a sample of the annotations, and having a discussion to see if their questions or comments reflect that engagement.

The scan of the reading as a whole is to see if the student has completed the reading and done a “reasonable” amount of annotation. (Each instructor will have their own sense of what amount is “reasonable,” and should convey that to the class.)

The close look at a sample is to see that there is a mix of content and reflection, as described above, and also to get a sense of the level on which the student is engaging with the text.

For DeLong (2022) mentioned above in footnote 1, we read the 18 chapters in three sets of six chapters, and each set of chapters was given two weeks (four class periods). The first of those four periods was replaced with small-group meetings, and each of the following three periods covered two chapters.

I organized the students into discussion groups of mostly three students, though some ended up as two.² Within each group, I quickly assessed each student's annotation, as described above, and noted their score. We then touched on highlights, confusing points, connections to other things in or outside of the course—there was no shortage of discussion material.³

In addition to DeLong (2022), each student, in consultation with me, chose a scholarly article to understand and present to the class. Once the article was chosen, the student had three things to accomplish in the next step of the project:

1. Annotate the article;
2. Use their annotations to prepare an outline of the article's contents;
3. Write a paragraph discussing some issue raised in the article that the student would like to know more about.

Since each student had a different article, the discussion and assessment of this work was one-on-one, rather than in small groups. The discussion was to address questions or misunderstandings that surfaced in the process of annotation, and to get a sense of the student's grasp of the paper's content. The outline was assessed for whether it plausibly grew out of the student's annotations (as opposed to having been generated by an LLM).

The paragraph led to a discussion of how to investigate the “side quest” that it identified. While most of the course structure was designed specifically to fence out the use of LLM's, this is an area where LLM use was allowed, though not required. My understanding of the current state of play with LLM's is that, in the hands of a person with some relevant knowledge, they can be time-savers, while they are highly unreliable tools for information about unfamiliar subjects. If a student is using LLM's for all of the course work, they are likely getting a substantial amount of inaccurate information while undercutting their own

²Rather than a sign-up, I created a survey with multiple time slots when I was available and asked students to provide information about their availability, with possible responses of, “Preferred,” “Available,” “Prefer not, but possible,” and “Not available.” I then used that information to create a schedule that got everyone in and tried to have everyone in a “Preferred” or “Available” slot. If I had done signups, people with broad availability might end up filling up slots that happened to be the only ones that worked for other students with more constrained schedules.

³I planned on 30 minutes for each discussion, but it turned out to be hard to keep myself strictly to that schedule.

learning. If they are looking further into a particular subject related to a paper that they have read and truly understood, in the context of a course where they are truly learning and reading, they are in a relatively good position to use an LLM constructively, with a nose for where it might be getting things wrong.

3.3 Written reflection

While I scheduled 30 minutes for each discussion, I asked the students to provide their availability for an 80-minute block of time (the same length as the class period that had been cancelled to give them and me for these meetings), in order to accommodate the composition of paragraphs, one for each chapter.⁴ In writing those paragraphs, students had access to their annotated books and were *expected* to be making use of that, in addition to drawing on the discussion we had just had if there had been something there that was useful for them.

As mentioned above, I modified the instructions for the paragraph over the course of the semester. The first version asked for a kind of summary of the chapter, and my intent there was to prod the students into “chunking”. But reading over what they produced, I realized that a one-paragraph summary of a 20- or 30-page chapter often leads to a sequence of sentences about different things, rather than a paragraph that is *about something*.

The revised assignment was to choose some significant element of the chapter, identify it, and discuss it. The goal was no longer to have them create an elevator-pitch version of the chapter, but to have at least one element of the chapter’s contents that they had understood well enough to write about. Rather than aiming for, “Know this entire book by having written good summaries of each chapter,” it moved in the direction of, “Know many important things from this book.” I later realized that if I *had* insisted on condensed chapter overviews, rather than reflection on particular points, I would have incentivized the use of an LLM to create that paragraph, with the student either rewriting it on site from partial memory or even tucking a printed version into their pocket, which they would then rewrite by hand after our discussion.

3.4 Accessibility accommodations

In my cohort of 20 students, none had physical obstacles to completing the assignments as designed. However, there are students for whom the structure described here would be either impossible or an unreasonable burden. The most obvious to me are:

- Visual impairments, where the student reads text in Braille, with an audio reader, or magnified on a device screen.

⁴The students swapped 80 minutes in the classroom for 80 minutes in small-group discussion and individual paragraph writing. I swapped an 80-minute class for 5 hours of discussions. I was fine with that trade, because I also saved the prep time that the conventional class period would have required.

- Physical inability to write by hand, whether due to neurological issues or the absence of digits or a hand.

If presented with such a case, I would not try to craft a solution myself, but would work with the student and the college's support services to find a useful way forward. I would be clear about my objectives with the assignments, and then see what approach we could create that would come as close as possible to meeting that while being reasonable for the student.

Though I didn't have any accessibility issues revealed to me among the 20 students in the class, I did have two students who had bought the book in electronic form before I had made clear the requirement to have a physical copy, leaving them with nothing to physically write on. Several other students went with the understandably popular option of renting the book, which meant they had a physical copy but didn't want to write in it because of the impact on its value at the end of the semester.

The students with e-books showed me their platforms, both of which supported a reasonable electronic version of writing on the page. Throughout the semester, their electronic annotations looked like work they would have produced, and their ability to discuss the text supported the impression that they were indeed the ones creating the annotations they were showing me.

Students with rented books were advised to have a stack of post-its by their side. The words of summarization or reflection that would have gone in the margin could instead go on a post-it, which would then be placed on the page that provoked the response.

4 Results

This was not constructed as an experiment with controls, so such data as I have are essentially anecdotal, but they are suggestive of success.

4.1 Student feedback

The epigraph at the start of this article speaks to the impact that the structure had on that particular student, getting a self-identified non-reader to knuckle down and do the work. Elsewhere in that same assignment, the student wrote, "While the debate on whether feudalism can be applied to India's early economic history isn't the most exciting theme, the sheer amount of work and nit-picking that went into it will have it logged in my memory."

In the end-of-semester student evaluations of the course, six of the 20 students gave written responses (18 students completed the numerical portion of the evaluation form). The six comments were all favorable to the course, with one that spoke directly to the pedagogical approach: "Coming into the course, I was skeptical about what the focus of the course would entail. The way that the

course was structured enabled my own critical thinking more than the average course.”

That same student did have a piece of constructive advice: “The only thing I would change is making the out-of-class Book Review and writing more frequently to help with time management from the students’ end.” It was interesting to glimpse that from the student side of things. I had put our first two-week set of six chapters in weeks 4 and 5, with less-demanding assignments (including practicing annotation) in those first three weeks, so that students would have time to read ahead. And I wasn’t coy about it either, giving weekly reminders that they were to have the first 180 pages annotated by the start of week 4; 10 pages a day, from two-and-a-half weeks out would get you over the line without breaking a sweat. But it was evident that many of the students had trouble motivating themselves to follow that relaxed schedule.

Another student’s comment ended, “overall was a good class and a good professor”; sadly, that positive note followed a “but,” and the thing before the “but” was, “For a Senior, involved in several positions and classes, it was too much ‘homework’ outside of the classroom.” Yet I was asking for less out-of-class work than what was expected of me as an undergrad in upper-level courses between 1985 and 1989, and was in line with my own expectations in earlier courses I’ve taught.

We can and should teach in ways that use students’ time well, towards defensible goals. But if we’re doing a reasonably good job on that front, the only way for students to spend less time on their course work is for them to learn less. This shows up in one of the common responses to the eruption of LLMs, which is to move all graded work into the classroom. The change is understandable, given the difficulty of knowing what written work produced out of our sight is actually by the student, but it also sacrifices the time spent individually wrestling with course material outside of the classroom, where a lot of learning used to take place.

My hunch is that this particular student, like many others, was used to being able to pass quickly through a reading, say something intelligent about it, and have done their work. The student cited earlier wanted help with time management, whereas this one wanted to simply spend less time on the course. A pedagogy based on annotation allows the instructor to get back some out-of-class time for course-learning purposes, which is one of the things that makes it valuable. But we should expect to have significant numbers of students who react to it as being “too much homework.”

4.2 Colleagues’ informal observations

Our small department has our offices clustered around a common space with a table where students study together or have conversations with faculty. We often have our office doors open, making us aware of the activity out at the table.

The small-group discussions on a set of book chapters happened in my office, but the follow-up writing happened at the table, often including discussion among the students who were there. And my one-on-one discussions regarding students' individual articles often happened at the table.

Colleagues repeatedly commented on the level of engagement and interest they observed in the students during these activities. They also reported student comments to them about how much they were getting from the class.

4.3 Quality of student assignments

The individual paragraphs were of mixed quality. Three students repeatedly struggled to write substantive responses; others wrote reasonably, but on minor issues; and some students fully grasped important parts of the text and discussed them well.

But the goal was to have students engaged with the reading, and the combination of that initial effort, feedback on their paragraphs, and class discussion of the readings seems to have worked for almost everyone, as suggested by their performance on the final exam.

The exam had two questions, available five days ahead of the exam period, with an element of choice in each question. Students were instructed to begin their answer by identifying what they had chosen, then spend three or more paragraphs explaining or discussing the subject of their answer. To end each answer, their assignment was to connect their topic to its possible use in their own future (question 1) or to something outside the course (question 2)⁵. They had guidelines of minimum numbers of different pieces and types of course-based material to cite in answering the questions.

They were allowed to bring a single 3x5 index card with notes, which could include references to the course materials they wanted to cite, brief outlines, or significant points. Beyond that index card, they had to write from what was in their heads.

17 of the 20 students took the final exam. Of those, 16 discussed the material at a level I considered "C-" or higher: some relevant thought, clear derivation from course materials, though possibly with inadequate citation, basically competent writing. Three students were in the "B" range, and eight in the "A" range.

The last part of each of the two final questions asked the students to build on the thing they'd written about and connect it to their future selves or to something

⁵The epigraph is taken from a response to question 1, which asked students to identify and discuss a thing from the course that would stick with them over a long period of time. The intent was for them to choose an idea from the course (e.g., Malthusian dynamics; the history of the corporation; the difference between Smithian growth and Schumpeterian growth). From that perspective, the student missed the point of the assignment, focusing on the method rather than the content. But it does speak to the impact that the technique had on him.

outside the course. Again, one of the 17 students struggled, but the other 16 all wrote intelligently and substantively.

10 of the 20 students presented their individual articles to the class. I have done similar assignments in the past, and even strong students are prone to misrepresenting some aspect of the work they’re discussing, or revealing some significant gap in their understanding of it. Of the 10 presentations, some had second-order issues of content but were basically strong, and others were excellent; none showed major gaps in understanding.

5 Why it works

The first two parts of this section explain why I think it’s important to limit AI use as part of my deeper goal of fostering long-lasting knowledge and understanding. The remaining five explain why I think annotation specifically, as described in this paper, is effective.

5.1 The importance of on-board knowledge

There is nothing abstract about thinking. It is the act of putting two or more things in relation to each other in your mind. This can be noticing what similarities and differences there are between two things. It can be deciding whether a certain theory and a certain piece of evidence are consistent. All of this implies that the ability to think and the ability to understand a “text” (in the sense of a thing we are reading, watching, looking at, or listening to) is conditioned on what we have in our minds when we encounter the text.

Similarly, reading is not an isolated, Platonic “skill” existing on an island unconnected to the reader’s knowledge of some relevant part of the world. This is demonstrated for K-2 students by Cabell and Hwang (2020).

In discussions of education more generally, critical thinking and problem-solving are skills that routinely get lip service, and sometimes are given actual support. Like reading, these skills depend on prior knowledge, as does the process of acquiring knowledge itself:

What you know determines what you see (Kirschner, 1991). Another and possibly more prosaic way of saying this is: Knowledge begets knowing. The more extensive one’s knowledge-base is in terms of both its breadth and its depth the more easily new knowledge is acquired and remembered (Alexander et al., 1994; Ausubel, 1968; Shapiro, 2004). Knowledge is also essential for carrying out the complex cognitive skills such as critical thinking (you think critically about something) problem-solving (you solve problems in something) and reading comprehension (you comprehend something written about something). The more robust one’s knowledge-base the more seamlessly and efficiently these complex cognitive skills—which

require students to ‘think deeply’ and are precisely those that teachers aim to develop in their students—are acquired and can be carried out.⁶

(Surma et al., 2025, p. 5, citations in the original)

And so your ability to think about a text, even to simply understand it, is rooted in what related knowledge you already have on board.

A possible rejoinder is that you can read things from two or three texts and think by making comparisons between/among them. The limitation on this is the difference between long-term memory and short-term memory, or working memory.

Long-term memory is functionally limitless—it’s not *literally* limitless, but putting things into long-term memory takes directed attention or many repetitions (Loewenstein and Wojtowicz, 2025, p. 12), with the result that the limit on the capacity of our long-term memory is not the “storage” space but the time and attention needed to put things there. “[I]t could be argued that attention is the only factor restricting the acquisition of human capital in many circumstances: if people could instantaneously absorb and process information without bound, acquiring knowledge that has already been developed—such as calculus, Latin, or the contents of this paper—would be effectively costless.” (Loewenstein and Wojtowicz, 2025, p. 32)

Working memory is very small, capable of holding five to 10 things at a time. That puts a hard limit on the number of things you can include in your thinking if you’re relying on only what’s in your working memory.

The solution is that your working memory has access to your long-term memory. Some of the contents are weakly stored or haven’t been pulled up in a while, and so they may only respond to a strong signal from the working memory. Other pieces of information are more readily available. To varying degrees, your working memory has access to all of it.

5.2 AI and learning

The metaphor of a double-edged sword conveys the idea of a tool that has beneficial uses but can also cause harm. AI is double-edged in a different sense. One of its effects is that it increases the importance of higher-level thought and insight. Another is that it makes it harder to get good at those skills: by enabling us to produce a simulacrum of thought without having built the foundation on which such thought rests, it disincentivizes the process of learning to think.

⁶The same principle applies to other “texts.” A piece of music “sounds different” after you have listened attentively to many other pieces of music in that genre. A painting “looks different” after you have carefully examined numerous paintings. A piece of creative writing has a different, fuller meaning for you if you know the literary allusions the writer is making, or the social context of certain phrases or idioms, and so on.

Ethan Mollick, a strong AI booster, spots the problem in a recent New York *Times* panel:

What I really worry about with junior workers is that field experience is often crucial to evaluating work you didn't create yourself, whether it comes from humans or A.I. ... If you have no experience, you can't do those things.

And that's what it takes to manage A.I.

Acemoglu et al. (2026a)

Mahroum (2026) nicely articulates the distinction between types of work that LLM's currently do well and others where we don't want to give them free rein: "While AI can handle tasks like reading, synthesis, and pattern recognition far more efficiently than humans and with relatively little risk, strategic judgment and interpretive reasoning are different. Those capacities must be deliberately maintained and strengthened, not quietly handed over to machines."

But if AI makes it easier to avoid building and maintaining the foundation of those higher-level skills, how do we protect them?

This quandary is explored in a formal economic model by Acemoglu et al. (2026b), where good decisions result from the combination of general knowledge and context-specific knowledge. Humans naturally pursue context-specific knowledge in order to solve particular problems facing them; in looking for solutions, they draw on the commonly-held stock of general knowledge and combine it with the context-specific knowledge that they have gathered. Crucially, while gathering context-specific knowledge, they end up contributing back to the store of general knowledge that is then available for people making decisions in the future. In their model, there are plausible states of the world in which humans stop contributing to the stock of general knowledge, leading to what the authors call a state of "knowledge collapse."

A big determinant of that risk is the accuracy of AI's responses: the more reliable they become, the less incentive humans have to spend time pursuing context-specific knowledge with its socially useful byproduct of general knowledge, which suggests a possible policy solution: "We show that limiting the effective precision of agentic recommendations via deliberate garbling can preserve learning incentives and prevent knowledge collapse." (Acemoglu et al., 2026b, p. 4) Beyond the irony of preserving the creation of knowledge by making knowledge harder to find, there's the low probability of getting either the sellers of AI or the users on board with something like this.

The act of relinquishing mental work as modeled in Acemoglu et al. (2026b) is not a merely theoretical possibility:

Another core aspect often overlooked in discussions of AI-driven learning is the value of background knowledge. When learners are exposed to the context or history behind a topic—even briefly—it

tends to stimulate curiosity, deepen engagement, and create cognitive hooks for long-term retention. In contrast, relying on AI to deliver direct answers bypasses this natural entry point into learning. Without understanding the ‘why’ behind a concept, learners may struggle to apply knowledge meaningfully or retain it over time. This undermines not only the joy of learning but also its sustainability and relevance in real-world settings.

(Hasan, 2025, p. 4693)

This phenomenon makes its way up the educational ladder, until you have a 10-year Princeton professor describing the increasing use of AI not just by students, but by faculty, in applications from creating and grading assignments to writing reviews of articles submitted for publication in peer-reviewed journals. This leads to what he calls “the shell university, where AI generated work is being passed off as human, and then in turn evaluated by another AI, again passed off as human. . . We are entering a new phase in universities where degrees will be handed out, and it may well not be possible to know if the *student did any work at all*.” (Truex, 2026, emphasis in original)

In a recent interview, the economist Philippe Aghion drew a policy recommendation from the way that AI increases the importance of thinking while making it harder to learn how to think. He praises Denmark’s “flexicurity” system of easy hiring and firing for economic flexibility, combined with *very* generous unemployment benefits to provide security and to spread the benefits of economic flexibility more broadly across society. “I believe all countries should have such a system, especially in the age of AI. But this model works best when it is backed by a good education system that teaches people how to learn. We need schools without AI where students learn to read, write, concentrate on a book from beginning to end, and perform calculations on their own. Those skills matter because they teach people how to adapt.” (Aghion and Johnson, 2026)

To sum up, I want to see my students having and expressing interesting thoughts. I also want to see them increasing the contents of their long-term memories, in support of the interesting thoughts I’m going to ask them to have, but also in support of future thoughts that I lack the imagination to even prompt them toward. Because of the way that AI interacts with the processes of knowledge acquisition and mental-skill development, there are times when machines need to be boxed out of students’ class experience. The method described in the next section aims to do this by focusing as directly as possible on the process happening in the students’ minds.

5.3 Observable process

When you tell someone to dig a ditch or cut some wood, you can physically model the process for them, and you can also watch them doing the task in order to possibly make observations about how they could do it better. And the person doing the task knows that they are making progress with it, even

if they're getting it done less quickly or effectively than their more experienced instructor.

When we ask students to “learn” a thing, or to “read and understand” a text, we generally don't have anything analogous. They can't watch us “learning” a thing or deepening our own understanding of it. We can't observe them in the performance of the task. And while they may be looking at slides from class, or passing their eyes across a passage of assigned reading, many of them don't know whether they are learning; as for understanding, some of them know they're not “getting it,” while some others think they are but then reveal that they didn't get it once we start asking them questions.

Annotation is a thing the teacher can model: I can share my own pages of annotated reading so that students can see the density of responses, the short phrases brutally summarizing important content, the longer comments reflecting a thought of my own that occurred in the course of reading. And while I don't watch the students do the work of annotation, I do look at the result and discuss it with them, which brings me as close as I can get to watching them learn, observing them engage in the process of reading for understanding.

5.4 The battle for attention

Long before the internet, people in wealthy countries had so much information available that the binding constraint on their ability to learn about the world was not their access to information, but the amount of attention they were able to devote to a topic and focus on it. Attention became the scarce resource in the learning process.

As discussed in Sec. 5.1, good thinking is built on having a large amount of relevant information stored in your long-term memory, and things move into long-term memory through the use of scarce attention.

Like learning and understanding, it's easy to tell students to pay attention, but hard to model the desired behavior. Annotation helps here because it is an impossible task if you're not paying attention. You can't highlight a particular passage *and explain why you've highlighted it* if you haven't been actually taking in the words on the page. In my class, the students' awareness of the requirement to annotate seems to have kept their attention on the page, in ways that then showed up in class discussions and on written assignments.

5.5 Questions for reading are ‘teach to the test’

Early in my teaching career, I would often assign a reading accompanied by questions for the students to answer. The idea was that students who could answer those questions would necessarily have absorbed some of the important ideas from the text.

I observed that the results were analogous to when we “teach to the test.” The

intent was that students would read the piece as a whole, gain some understanding of it, and answer the questions from that place of understanding. In reality, a common outcome was that students would read superficially, looking for the specific material that addressed the specific questions I had asked. The students weren't engaging with the text but were instead hunting for the requested information and filtering out the rest.

To get away from that, I started requiring students to submit outlines of assigned readings. That was much more effective in getting them to read holistically and then go back and observe specific content. But the release of LLM's made that no longer a useful assignment, because the machine is now very effective at summarizing a text into a structured outline. The results aren't usually inspired or insightful, but they would generally warrant at least a "B" if they were actually done by a student.⁷

In annotating, as with making a good outline, you are engaging with the text in an authentic way, simply observing what's there. As you respond to the text, you are making your own meaning. And as you read, pen in hand, you are having to decide what things within the text warrant more of your precious focus; you are learning to allocate your attention in constructive ways. This is crucial because, "People can only learn from information they collect, process, and remember. As a result, attentional mechanisms are central to what people do (and do not) learn." (Loewenstein and Wojtowicz, 2025, p. 31)

5.6 Process, not product

The set of activities described in Sec. 3 results in various "products" (the words students have written in the margins of the text, the response paragraphs they write after a discussion or a lecture), but the annotations themselves are the direct emanations of the process that I'm really aiming at. The importance of process is articulated by the poem "Process-not-product" (Wilkinson, 2025)⁸:

Process rules, master of product.
The product serves the master.
When I follow process, the product is certain.
Process-not-product.
Process-not-product.
The only way forward.

In a related vein, Seybold (2026) talks about a "false binary between process and product." In his framing, the product we care about is the production of collective knowledge in a social environment, but since that is hard to measure, we expose students to repeated experiences of reading texts, listening to professors, and having discussions. "And then at the end of the term when you

⁷An LLM could probably do an equally creditable job with my earlier activity of reading questions, but I haven't tried that out.

⁸Written at the request of composer Michael Markowski, who then set it and others by Wilkinson in a collection of "Affirmations"

write me a close reading of the red wheelbarrow, or you do your exam, that will be sufficient evidence that you have been getting the product that is really the thing of value. And what AI has definitely done is eradicate the reliability of the relationship between the real product and the product we use as evidence of the existence of the real product.” (Seybold, 2026, 1:17:30 to 1:19:10; edited for clarity)

I would reframe Seybold’s point. He sees a superficial product (the exam answer or finished essay) and a real product (“the production of collective knowledge in a social context”). But note that the *production* of knowledge is a process, not a product. I agree with Seybold that AI has severed the relationship between the products we ask students to hand in and the process of learning that we ultimately care about.

5.7 Why not just take notes?

There is nothing new about the practice of taking notes on a reading, just as (some) students have long taken notes on a professor’s lecture. So why annotate? Why write directly into or onto a text, or require students to do the same?

First, there is the efficacy of writing by hand. As an example, Mangen et al. (2015) show greater word recall of hand-written words, compared to words typed on either a conventional keyboard or a tablet screen.

In Marano et al. (2025), we see some of the mechanisms involved:

The comparison between handwriting and typing reveals important differences in their neural and cognitive impacts. Handwriting activates a broader network of brain regions involved in motor, sensory, and cognitive processing, contributing to deeper learning, enhanced memory retention, and more effective engagement with written material. Typing, while more efficient and automated, engages fewer neural circuits, resulting in more passive cognitive engagement. These findings suggest that despite the advantages of typing in terms of speed and convenience, handwriting remains an important tool for learning and memory retention, particularly in educational contexts.

(Marano et al., 2025, p. 23 of 27)

Interestingly, given that evidence of handwriting supporting memory and learning by engaging more of the brain, Shibata and Omura (2018) find that handwriting outperforms typing because it is less cognitively demanding—that is, it uses less of the brain’s very limited working memory, presumably leaving more for the targeted subject matter. On the other hand, Stadler et al. (2024) find that use of LLMs interferes with learning, because it reduces the cognitive load. It may be that handwriting allows the working memory to focus on what the person is trying to learn, while LLMs allow the learner to avoid engaging the working memory on anything, including the targeted knowledge.

Effects of different note-taking modalities are documented in Mueller and Oppenheimer (2014), in which students who took notes in longhand with pen on paper did better on conceptual questions related to the material. It seems the pen-on-paper note-takers were doing less verbatim transcription and were more likely to process information and reframe it into their own terms.

All that said, even if we accept the relative value of writing notes by hand, that can be done in a notebook students keep next to them as they read. Why not simply have them show you their hand-written notes?

Part of the answer is an essentially “policing” function. A student could in principle ask an LLM to produce notes on the assigned reading and then copy out those notes by hand in order to complete the assignment. That’s more work than just printing out the machine’s response and bringing it to class, but it’s less work than actually doing the reading and creating the notes yourself.⁹ In contrast, we are still some steps of robot development away from a machine that can write an appropriate remark, by hand, at the appropriate place in a physical text. When I see the student’s own handwriting directly on the physical text, I have the highest possible confidence that they have done the work themselves.

Moving from policing to the substance of learning, there is some evidence that the neural networks for navigating the world are integrated with those for episodic memory, so the act of remembering is linked with location (e.g., Chandra et al. (2025)). In the literature, this is about the place where you are, such as a room. I speculate that these mechanisms of correlating locations and ideas may also apply to a text. When I write on a text, I am hooking my words to the physical location within the book that inspired those words. There may also be a relationship to the “method of loci,” colloquially known as a “memory palace” (extensively documented in a review article by Ondřej (2025)).

Additionally, annotation on the page brings you immediately to the larger context of the part of the text that provoked a particular thought. When you find a note that says, “reasons for declining trust in institutions,” you have the reasons on the page in front of you, but you also have the passage leading up to that, in case you need a reminder of the surrounding material that motivated the author’s introduction of those reasons.

6 Conclusion

In the wake of AI erupting into everyday life, teachers who want to be effective have had to think carefully about our goals and about how to reach them when

⁹It’s true that, in copying out the notes, the student might accidentally learn something, simply by the relatively slow passage of the words from the screen, through the student’s mind (briefly) and onto the page of the notebook. But if the distillation were enough for deep learning, we would get rid of books altogether and just have students read the distillations. The physical act of writing a note can reinforce the learning that happens as you do the work of creating the note in the first place, but a student who merely copies out notes created by a machine or a classmate has none of that earlier stage to reinforce.

our students have ready access to machines that “think.” Annotation on the printed page is a promising process for building the knowledge base and critical thinking that will enable students to engage fruitfully with AI in other contexts, where such engagement may be either unavoidable or desirable. It moves focus off of the product that machines can now spit out and onto the most granular, foundational step of the process that is behind the creation of such a product when a human is doing it. In a small, initial experiment, it has resulted in work of higher quality than I have seen in years.

I don’t believe in silver bullets, and this isn’t one—not every student in the class succeeded. But most showed signs of wanting to learn, and this structured approach to annotation gave them a clear path toward that goal.

The old saying cautions us that, while you can lead a horse to water, you can’t make it drink.¹⁰ But if the horse is thirsty, should you leave it to die as the grasslands around it dry up? If our students are thirsty and we lead them to water, what do we think they’re going to do?

References

- Acemoglu, D., Ball, D., Mollick, E., Shih, C., and Wasik, B. W. (2026a). Who will actually thrive in the hybrid A.I.-human work force: A panel of experts explains how job seekers should prepare for the future of work. *New York Times*.
- Acemoglu, D., Kong, D., and Ozdaglar, A. (2026b). AI, human cognition and knowledge collapse. Technical report, National Bureau of Economic Research.
- Aghion, P. and Johnson, S. (2026). Why is europe’s economy falling short?
- Cabell, S. Q. and Hwang, H. (2020). Building content knowledge to boost comprehension in the primary grades. *Reading Research Quarterly*, 55:S99–S107.
- Chandra, S., Sharma, S., Chaudhuri, R., and Fiete, I. (2025). Episodic and associative memory from spatial scaffolds in the hippocampus. *Nature*, 638(8051):739–751.
- DeLong, J. B. (2022). *Slouching towards Utopia: An economic history of the twentieth century*. Hachette UK.
- Hasan, M. K. (2025). How AI quietly undermines the joy and effort of learning: a call for rebalancing education in the digital age.
- Loewenstein, G. and Wojtowicz, Z. (2025). The economics of attention. *Journal of Economic Literature*, 63(3):1038–1089.

¹⁰Or in a cynical turn, you can lead students to Descartes, but you can’t make them think.

- Mahroum, S. (2026). Cognition for sale. *Project Syndicate*.
url=<https://www.project-syndicate.org/onpoint/what-ai-cannot-replace-by-sami-mahroum-2026-06>.
- Mangen, A., Anda, L. G., Oxenburgh, G. H., and Brønnick, K. (2015). Handwriting versus keyboard writing: Effect on word recall. *Journal of writing research*, 7(2):227–247.
- Marano, G., Kotzalidis, G. D., Lisci, F. M., Anesini, M. B., Rossi, S., Barbonetti, S., Cangini, A., Ronsisvalle, A., Artuso, L., Falsini, C., et al. (2025). The neuroscience behind writing: Handwriting vs. typing—who wins the battle? *Life*, 15(3):345.
- Mueller, P. A. and Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. *Psychological science*, 25(6):1159–1168.
- Ondřej, J. (2025). The method of loci in the context of psychological research: A systematic review and meta-analysis. *British Journal of Psychology*, 116(4):930–986.
- Seybold, M. (2026). The secret history of Canvas LMS, corporate raiders, and the chatbot bubble (Vandal Live at UVU).
- Shibata, H. and Omura, K. (2018). Reconsideration of the effects of handwriting: Comparing cognitive load of handwriting and typing. *ITE Transactions on Media Technology and Applications*, 6(4):255–261.
- Stadler, M., Bannert, M., and Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior*, 160:108386.
- Surma, T., Vanhees, C., Wils, M., Nijlunsing, J., Crato, N., Hattie, J., Muijs, D., Rata, E., Wiliam, D., and Kirschner, P. A. (2025). How knowledge matters. In *Developing Curriculum for Deep Thinking: The Knowledge Revival*, pages 5–35. Springer.
- Truex, R. (2026). The shell university: AI has taken something fundamental from the business of teaching and learning. *The Chronicle of Higher Education*.
- Wilkinson, A. (2025). Process-not-product. In *Affirmations*. Markowski Creative (ASCAP).
- Wyman, P. (2025). Crixus, the Spartacus rebellion, and resistance to slavery. Perspectives: Past, Present, and Future, by Patrick Wyman.