By Owen Kichizo Terry

Look at any student academic-integrity policy, and you'll find the same message: Submit work that reflects your own thinking or face discipline. A year ago, this was just about the most common-sense rule on Earth. Today, it's laughably naïve.

There's a remarkable disconnect between how professors and administrators think students use generative AI on written work and how we actually use it. Many assume that if an essay is written with the help of ChatGPT, there will be some sort of evidence — it will have a distinctive "voice," it won't make very complex arguments, or it will be written in a way that AI-detection programs will pick up on. Those are dangerous misconceptions. In reality, it's very easy to use AI to do the lion's share of the thinking while still submitting work that looks like your own. Once that becomes clear, it follows that massive structural change will be needed if our colleges are going to keep training students to think critically.

The common fear among teachers is that AI is actually writing our essays for us, but that isn't what happens. You can hand ChatGPT a prompt and ask it for a finished product, but you'll probably get an essay with a very general claim, middle-school-level sentence structure, and half as many words as you wanted. The more effective, and increasingly popular, strategy is to have the AI walk you through the writing process step by step. You tell the algorithm what your topic is and ask for a central claim, then have it give you an outline to argue this claim. Depending on the topic, you might even be able to have it write each paragraph the outline calls for, one by one, then rewrite them yourself to make them flow better.

The common fear among teachers is that AI is actually writing our essays for us, but this isn't what happens.

As an example, I told ChatGPT, "I have to write a 6-page close reading of the Iliad. Give me some options for very specific thesis statements." (Just about every first-year student at my university has to write a paper resembling this one.) Here is one of its suggestions: "The gods in the Iliad are not just capricious beings who interfere in human affairs for their own amusement but also mirror the moral dilemmas and conflicts that the mortals face." It also listed nine other ideas, any one of which I would have felt comfortable arguing. Already, a major chunk of the thinking had been done for me. As any former student knows, one of the main challenges of writing an essay is just thinking through the subject matter and coming up with a strong, debatable claim. With one snap of the fingers and almost zero brain activity, I suddenly had one.

My job was now reduced to defending this claim. But ChatGPT can help here too! I asked it to outline the paper for me, and it did so in detail, providing a five-paragraph structure and instructions on how to write each one. For instance, for "Body Paragraph 1: The Gods as Moral Arbiters," the program wrote: "Introduce the concept of the gods as moral arbiters in the Iliad. Provide examples of how the gods act as judges of human behavior, punishing or rewarding individuals based on their actions. Analyze how the gods' judgments reflect the moral codes and values of ancient Greek society. Use specific passages from the text to support your analysis." All that was left now was for me to follow these instructions, and perhaps modify the structure a bit where I deemed the computer's reasoning flawed or lackluster.

The vital takeaway here is that it's simply impossible to catch students using this process, and that for them, writing is no longer much of an exercise in thinking. The problem isn't with a lack of AI-catching technology — even if we could definitively tell whether any given word was produced by ChatGPT, we still couldn't prevent cheating. The ideas on the paper can be computer-generated while the prose can be the student's own. No human or machine can read a paper like this and find the mark of artificial intelligence.

When we want students to learn how to think, assignments become essentially useless once AI gets involved.

There are two possible conclusions. One is that we should embrace the role AI is beginning to play in the writing process. "So what that essays are easier to write now? AI is here for good; students might as well learn to use it." Of course, it's important to learn to put together a cohesive piece of written work, so it makes perfect sense to embrace AI on assignments that are meant to teach this skill. In fact, it would be counterproductive not to: If a tool is useful and widely available, students should learn how to use it. But if this is our only takeaway, we neglect the essay's value as a method for practicing critical thinking. When we want students to learn how to think — something I'm sure all educators consider a top priority — assignments become essentially useless once AI gets involved.

So rather than fully embracing AI as a writing assistant, the reasonable conclusion is that there needs to be a split between assignments on which using AI is encouraged and assignments on which using AI can't possibly help. Colleges ought to prepare their students for the future, and AI literacy will certainly be important in ours. But AI isn't everything. If education systems are to continue teaching students how to think, they need to move away from the take-home essay as a means of doing this, and move on to AI-proof assignments like oral exams, in-class writing, or some new style of schoolwork better suited to the world of artificial intelligence.

As it stands right now, our systems don't accomplish either of those goals. We don't fully lean into AI and teach how to best use it, and we don't fully prohibit it to keep it from interfering with exercises in critical thinking. We're at an awkward middle ground where nobody knows what to do, where very few people in power even understand that something is wrong. At any given time, I can look around my classroom and find multiple people doing homework with the help of ChatGPT. We're not being forced to think anymore.

People worry that ChatGPT might "eventually" start rendering major institutions obsolete. It seems to me that it already has.

Owen Kichizo Terry is an undergraduate at Columbia University.

Will ChatGPT Change How Professors Assess Learning?

It won't be easy without their colleges' support.

BRAVE NEW WORLD

Quelle: https://www.chronicle.com/article/will-chatgpt-change-how-professors-assess-learning

By Beckie Supiano

APRIL 5, 2023

The worst-case scenario came into focus almost as soon as ChatGPT was released: Students could feed a professor's prompt into the chatbot, collect its response, turn it in as their own, and get credit without doing any work at all. Go through those motions enough times, one imagines, and a student could pass a class — even earn a degree — without the learning that those things are supposed to signify.

For three years now, professors have adjusted their teaching, and adjusted it again, to adapt to a global pandemic, remote instruction, and a student population whose actions and expectations have changed. What fresh new hell was this?

ChatGPT, which runs on a "large language model," a word predictor that has been trained with enormous amounts of data, has been greeted as a game changer in many domains. The potential uses of the tool, and others like it, extend far beyond the completion of college coursework. And many professors are excited by the technology's potential to enhance learning, and perhaps provide needed support to students who start at a disadvantage. There are lots of ways students could use ChatGPT without having it do their work for them, like using it to brainstorm ideas or offer clearer definition of something they're trying to understand.

But many professors are apprehensive. What does it mean if a text generator can complete assignments as well as an undergraduate? Will the advent of these generative artificial-intelligence systems force faculty members to change the way they assess student learning all over again?

Let's unpack that worst-case scenario a bit. Why are professors collecting students' responses to that prompt in the first place? When it comes to whether a student has learned a skill, a professor can get a pretty good idea by watching the student use it (though that does require some judgment), says Betsy Barre, executive director of the Center for the Advancement of Teaching at Wake Forest University. Figuring out whether students have acquired knowledge is harder, though, Barre says, "because we can't see inside your brain." So professors collect evidence of that learning — for instance, by having students answer questions designed to reveal their ability to retrieve and use particular knowledge.

Professors provide two kinds of assessment, summative and formative. Summative assessment, like the grade students get on a test or in a course, captures where students stand at a point in time. Formative assessment, like

the comments professors leave on the draft of a paper or a quiz meant to check students' understanding, is feedback meant to support learning by letting students know what they need to work on.

If students hand in assignments completed by ChatGPT, then those assignments can't give professors the information they need about students' learning.

This AI thing? Yeah, it's another challenge. But, OK: It's another challenge.

This introduces a couple of different potential problems, Barre says. One is whether professors can give meaningful summative assessments. If they can't, employers and graduate schools may not be able to rely on the signal of students' grades. That, Barre says, isn't really a college teaching problem.

But the scenario gums up formative assessment, too. If the work that's turned in wasn't completed by students, it can't reveal what they know or can do. And that is very much a college teaching problem.

Experts agree that generative AI is here to stay, like Wikipedia, the internet, the calculator, and all manner of innovations predating it. "You can't fight technology," says Adam Fontecchio, inaugural director of the Center for the Advancement of STEM Teaching and Learning Excellence and a professor of electrical and computer engineering at Drexel University. Fontecchio, who is part of a working group on artificial intelligence at Drexel, has told his own students they can use generative AI, but asks them to disclose its use. Students are using the tool to write sections of their reports for a group project, Fontecchio says, and some are using it to help design circuits and write code, too. "I'm aware of a team that is using it right now to try and design a circuit that is doing some things that aren't the normal way circuits work — they're trying to push voltage limits and current limits."

Part of what makes ChatGPT valuable, Fontecchio says, is that the more someone uses it, the better the tool gets at giving them what they want. Students are getting more personalized support than traditional online research could provide. And unlike the instructor of a course with 85 student teams to oversee, the tool is always available. "I don't know whether the circuit is going to turn out," Fontecchio says. But this team's project "feels way more advanced than anything I would have seen a team do in the past."

The project is meant to model for students nearly finished with college the kinds of work engineers do on the job. The students have plenty of incentives to do the work. Still, this example illustrates that ChatGPT can help students take risks, tackle something ambitious, and learn.

There's also a general consensus among learning experts that professors should make a point of talking about the tool, and their expectations for students' use of it, in their courses, and that the appropriate use of the tool in coursework will probably vary from discipline to discipline.

If the concern is that students could cheat, it's worth remembering that they could cheat six months ago and 60 years ago. Students taking a brand-new exam could already get answers to test questions in minutes from services like Chegg. Students could already plagiarize — or pay someone else to write their entire paper. With the entrance of ChatGPT, "what's changed is the ease and the scope," says Kevin Gannon, director of the Center for the Advancement of Faculty Excellence at Queens University of Charlotte, and a frequent contributor to *The Chronicle*.

It's unclear how well instructors will be able to spot students using AI in defiance of course policies. But experts discourage the policing of cheating anyhow, and most of the steps they encourage to mitigate cheating

still apply here. Professors can create conditions in which cheating is difficult, giving closed-book, closed-note, closed-internet exams in a controlled environment. They can create assignments in which cheating is difficult, by asking students to draw on what was said in class and to reflect on their own learning. They can make cheating less relevant, by letting students collaborate and use any resource at their disposal.

Or they can diminish the forces that make cheating appealing: They can reduce pressure by having more-frequent, lower-stakes assessments and increase students' desire to do their own work by making assessment more like work students might do after college, more meaningful, and maybe even enjoyable.

Ceceilia Parnther is an education scholar whose research focus is academic integrity. While the phrase "academic integrity" is often used as an antonym for cheating, experts like Parnther mean something more expansive by it. She describes her interest as "understanding how students know what they know, and how they understand how to learn."

Parnther's approach to ChatGPT in her own courses is to look at the questions she poses to students and ask herself what is lost if they feed the question into some kind of bot. From there, says Parnther, an assistant professor of higher-education leadership at St. John's University (New York), "I have two choices. I can either change that prompt, or I can show students how to critically engage with the prompt such that they understand that they need and deserve more than what's coming out." If there's a prompt that seems important, but too easily completed by AI, it can help to adjust it so that a question will "require someone to use their personal experience, or their passion, or their creativity," Parnther says. It's also important, she adds, to help students understand how ChatGPT works and teach them to interrogate the information it provides.

It hasn't been long since remote instruction pushed professors to reconsider their assessments. How well have the changes they made then teed them up for this new set of challenges?

ChatGPT poses little risk to an instructor who's giving a traditional exam in a tightly controlled environment, says Mike Caulfield, a research scientist at the University of Washington's Center for an Informed Public. But many instructors moved away from that during the pandemic, because they could not or did not want to try to create such testing conditions online.

For instance, professors who gave multiple-choice exams before might have adjusted those tests to include more short-answer questions. They might have given students more time, allowed them to consult more resources. Those short-answer questions are now the kind of thing a student could easily use ChatGPT to complete. So Caulfield suspects that professors who made an effort to improve their tests now find themselves back at the drawing board.

Jessica Bickel sees it differently. "We've already had to go through the grinder once to figure out how in the world do I adapt," says Bickel, who is serving as a facilitator in a faculty learning community on AI at Cleveland State University, where she is an associate professor of physics. "This AI thing? Yeah, it's another challenge. But, OK: It's another challenge." If professors were able to figure out emergency remote instruction, she believes, they can figure out whatever comes next.

In the ideal scenario, ChatGPT could push the whole teaching-and-learning enterprise in good directions. Johanna Inman, who directs the teaching center at Drexel and serves on the AI working group, can imagine it. The disruption caused by AI, Inman hopes, "will move us even further."

That shift, in Inman's view, has two main strands. One is the way professors use class time. The ability to tape lectures has already freed class time up for other activities, she points out. If professors are worried that students will overrely on ChatGPT to complete their out-of-class assignments, professors may decide to use class time for whatever they most want to ensure students are able to do. The time could be devoted to practice,

research, and problem solving, she says. "It's that skill-based, collaborative learning that we know is going to benefit them."

The other strand is authentic assessment. Professors increasingly see the value in asking students to use what they know to solve real-world problems. ChatGPT and similar tools exist in that real world, and students will use them, in college and after. "We need to help students learn how to use the tools that they have at their fingertips in professional, ethical, disciplinary-based ways."

There's more going on here than the possibility that students could cheat more easily. If ChatGPT can generate text that plausibly resembles student work, what does that say about the work that's being assigned?

Some will argue that an assignment ChatGPT can excel in must not be a very good assignment. But there are plenty of other instances in which professors ask students to do something by hand that could be done faster using technology. The trick is knowing which tasks develop foundational knowledge that students will draw on to do more sophisticated work, and which are just grunt work we should all be grateful to give up. Do students need to know how to build a graph if an AI tool can make it instantly? Or basic coding? As technology advances, professors have to keep asking themselves which skills really matter.

Perhaps ChatGPT could help surface hidden talent.

John Warner's in the camp of professors who see AI's ability to ace writing assignments as an indictment of those assignments. A longstanding critic of the way in which writing is generally taught, Warner, the co-founder of a teaching-focused higher education [consulting firm,](#) has created a [self-paced online course](#) to help instructors think through the implications of ChatGPT and similar tools.

The actions Warner encourages to ensure that writing assignments can't simply be farmed out to ChatGPT are ones he thinks are a good idea in the first place. Professors can emphasize reflection, and also the writing process. They might grade writing on a single-point rubric, like whether a piece has "energy" — a quality that some professional writers like Warner find missing from AI generated text. ChatGPT, Warner reminds us, can't think and doesn't read. It can put on a voice, but doesn't have its own.

At the same time, Warner says, the tool may push professors to reconsider the way writing is used for assessment in courses that are not designed to develop writing. "One of the things that may come up is we discover that certain written assessments we've done because we think students should be writing could be better done in other ways," he says.

The arrival of ChatGPT has often been compared with the arrival of the calculator: Here's a powerful, automated way to instantly get something that used to take more effort. But there's an important difference. As long as a student puts the right equation into the calculator, the answer it generates will be correct. Not so ChatGPT. Others have likened the tool to early-days Wikipedia, which was a great resource on topics in which one was already well informed. One big question is whether it becomes more like current Wikipedia — for most of us on most topics, a great starting point.

The conversation about ChatGPT, like the remote-instruction conversation that came before it, highlights an ever-present teaching challenge: There is no perfect way to assess learning. No matter how well a test is designed, students with good test-taking strategies who perform well under pressure will score higher than classmates with similar understanding of the material who don't have those advantages. The same is true on writing assignments. A student who's a strong writer is likely to get a higher grade than a peer with a similar grasp of the material whose writing skills are less developed. Should that be the case, though, in a course that doesn't teach writing?

Writing can be a path for thinking, Caulfield says, and if ChatGPT weakens that connection for students, it will be a loss. But when it comes to assessment, the relationship between good writing and good thinking is not clear cut. There are students whose fluid prose obscures an undercooked thesis. And there are students who have good insights but struggle to convey them on the page in the desired style. Perhaps, he thinks, ChatGPT could help surface hidden talent.

So where does that leave us? Professors can structure the way students write in their courses to emphasize process, not just product. They can read for substance, and not just style. Faculty members can ask students to reflect, to bring something of themselves into their assignments. They can explain that education is more than an exchange of tuition dollars for a diploma; show them the real value in learning the things they want students to learn, even when doing so is unpleasant. They can help students summon the motivation and even excitement that make the idea of using a text-generator to get a decent grade for minimal effort unappealing.

Or professors *could* do all of those things, if they worked within a system designed to prioritize good teaching.

"ChatGPT in many ways is the biggest chicken that's come home to roost," says Gannon, the Queens University teaching-center director. "You can't have a conversation as an institution about 'how do we support faculty to make sure that their classes are done with integrity' if you're not talking about supporting adjunct faculty and not teaching classes with 200 students."

For a while there, it seemed like the circumstances of the pandemic — which, for a time, collapsed the whole student experience into the classroom, hastily reconstructed in Zoom, and saw professors holding it all together with gum and string — might lead to a reckoning about the ways in which higher ed's status quo devalues teaching. Certainly teaching centers ramped up their work in the pandemic, and professors probably spent more time on their teaching than ever before. But the industry appears to have dodged any real pressure to adjust its model of providing instructors with minimal preparation or support and then loading them up with ever-increasing demands.

Will colleges' response to this disruption be any different?