



A Command and General Staff College student uses ChatGPT to do a school assignment. (AI generated image by Michael Lopez, *Military Review*)

How to Think About Integrating Generative AI in Professional Military Education

Maj. Patrick Kelly, U.S. Army

Maj. Hannah Smith, U.S. Army

In November 2022, OpenAI's Chat Generative Pre-Trained Transformer (ChatGPT) became publicly available. The artificially intelligent chatbot is revolutionary. ChatGPT generates detailed and seamless text in response to user prompts. It can write essays and compose poetry; it can debug and write computer code. No previous technology has accomplished these tasks so quickly and effectively. And ChatGPT's release was but one highly public moment among many recent advances in available generative artificial intelligence (AI). In 2023, Microsoft released a new version of its Bing search engine powered by AI, and OpenAI unveiled its new GPT-4 model.¹

Since ChatGPT's release, generative AI has dominated headlines, especially within the academic community. Well-known publications featured arguments that "the college essay is dead" and no one is prepared for the transformation of academia, that ChatGPT is a "threat to higher education," and that higher education must "change or die."² The chatbot was even compared to a "plague," evoking images of COVID-19's devastating impact on education. Some schools responded by immediately banning access to ChatGPT on their networks.³ Others have looked to AI content detection software to identify and punish students using generative AI to complete their homework assignments.⁴

The U.S. Army is not immune to these developments. The Army profession requires officers to attend professional military education (PME) throughout their careers. PME educators, like their counterparts in civilian academic institutions, are struggling to negotiate generative AI. Part of the challenge stems from the character of ongoing debates about generative AI. Frequently, these debates are extreme—featuring polarized visions of higher education's death or revolutionary gains to productivity—and abstract, featuring arguments at the level of the technology itself rather than at the level of its specific use cases in learning. Debates at the conceptual level are critical, but they are not always helpful for instructors who must decide how to respond in the classroom right now.⁵

The Command and General Staff College (CGSC), which recently became an accredited institution issuing master's degrees to graduating students, has responded with a policy that allows individual courses to determine acceptable use of generative AI.⁶

However, in practice, the overwhelming message to CGSC students has been that they should avoid the technology altogether or risk losing their PME credit. The prohibitive implementation of the policy suggests that Army PME is highly sensitive to the risks of generative AI. There are indeed risks in adoption, especially uncritical adoption without consideration of potential costs. But there are also risks in general avoidance of the technology. As the Department of Defense's *Data Strategy* makes clear, Army leaders must be data-literate.⁷ Thus, the challenge for CGSC and other PMEs is to enable and encourage uses of AI which will augment students' learning while mitigating uses that will hamper students' development of critical-thinking skills.

This article frames two opposing views of generative AI within the context of Army PME: outright resistance of generative AI, on the one hand, and unquestioned acceptance of AI on the other. PME educators will benefit from a pragmatic and measured perspective that avoids these extremes and orients on the purpose of PME. Educators can simultaneously acknowledge the important benefits of AI while recognizing the need for guardrails on its use, since unconstrained adoption could undermine our mission to educate. Embracing this perspective and forgoing the extreme alternatives yields several practical recommendations outlined in this article.

Maj. Patrick Kelly, U.S. Army, is an assistant professor in international affairs in the Department of Social Sciences at the U.S. Military Academy. An infantry officer, Kelly has served in the 101st Airborne Division and 4th Infantry Division. He holds a PhD and an MA in political science from Stanford University and a BS in international relations from the U.S. Military Academy.

Maj. Hannah Smith, U.S. Army, is a student and Art of War Scholar at the Command General Staff College at Fort Leavenworth, Kansas. A military intelligence officer, she previously served as an assistant professor of international relations in the Department of Social Sciences at the U.S. Military Academy, in the 82nd Airborne Division, and in the 25th Infantry Division. She holds an MA in global affairs from Yale University and a BS in economics from the U.S. Military Academy.

Avoiding the Extremes

One possible reaction to generative AI is outright rejection of its use in PME. Instructors at CGSC may project this perspective to students by messaging that ChatGPT only leads to plagiarism cases and possible expulsion. PME leaders who adopt this view may be wary of AI technologies for a variety of reasons. Aware that ChatGPT can write complete essays, they may deny any valid uses, which might motivate them to prohibit student use of AI in any capacity. Worse, they may persist in using assignments that most incentivize the use of generative AI even as they outlaw such use. They may assume more rampant unauthorized use of generative AI than is warranted. And they may fail to attend to the technology's important and long-term implications or even to learn about its basic functionality.

Universal rejection is too dismissive of potential benefits to professional education and national security—two spheres that every graduating field grade officer occupies. Failure to leverage AI tools in any capacity will mean forfeiting competitive advantages for our students, institution, and national defense. Crude, complete bans of AI tools will undermine our mission in the name of preserving it. Guardrails on AI's use in the form of specific prohibitions in syllabi and policies for assignments meant to develop critical-thinking skills are useful, but instructors should avoid general hostility toward the technology and antagonism toward students who use it.

Another extreme reaction to ChatGPT's emergence is uncritical acceptance of generative AI as the latest calculator or word processor. In this view, generative AI is simply the latest technology that will improve higher education as soon as faculty and staff get on board. Uncritical proponents seek the unquestioned adoption of generative AI throughout education.

Chatbots are indeed like earlier disruptive technologies along a couple dimensions. One similarity is the initially resistant reaction of some educators, echoing debates (and even small protests) against the adoption of calculators in the 1980s and concerns about classroom use of computers in the 1990s.⁸

Another similarity is that chatbots, like earlier technologies, may enhance student learning by substituting for task completion and rule compliance (ChatGPT can spell check, write in complete sentences, and so forth), freeing up student energy for creative thought.

Advocates for generative AI in higher education are correct in noting the technology's potential for improving student learning. Generative AI can help students brainstorm interesting topics for a paper, act as a conversational tutor for general learning, or provide background context for more advanced topics. AI tools can also produce functional programming scripts, like for cleaning and analyzing data, with simple written instructions from the user—a capability that has some observers calling English the next big programming language.⁹ In applications like these, generative AI can support student learning by accelerating work and increasing productivity without compromising the learning process.

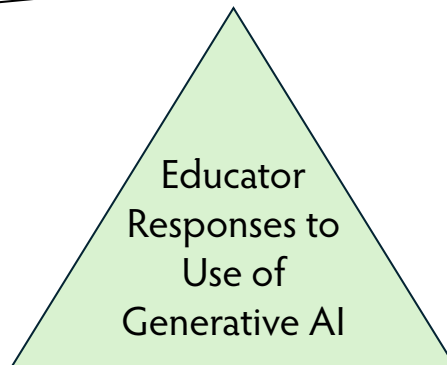
However, the problem with universal adoption is that generative AI can do much more. ChatGPT can synthesize evidence, generate a thesis statement, and develop a coherent and comprehensive argument to support it, all while balancing contradictory information. Within the context of CGSC, a student might upload doctrine files and ask ChatGPT to produce a summary of selected excerpts. And in these capacities, ChatGPT is clearly more than the latest typewriter. The differences are apparent when we attend to specific use cases and remind ourselves that knowledge is not only a product but also a process.¹⁰ Sometimes, ChatGPT can complement critical thinking by providing information and efficiencies. But other times, the technology can substitute for critical thinking.

Skeptics dismiss this point too quickly based on ChatGPT's known limitations, like tendencies to “confidently” misinform, “hallucinate” facts, or fail to cite sources. However, the key point in this debate is not whether generative AI can perform all written tasks better than humans, or whether a chatbot is indeed intelligent, but that it can consistently produce passable work, which creates incentives for students to shortcut processes of discovery and deep thinking.

Those who adopt the view of “uncritical acceptance” are too dismissive of tradeoffs. PME that implements this technology in every aspect of its program will certainly teach students useful skills, like how to efficiently interact with AI, but often at the expense of developing students' higher order thinking abilities. The importance of developing these abilities is explicitly embedded in CGSC's vision to develop field grade officers who can think critically

Outright
Rejection

Uncritical
Acceptance



(Figure by authors)

Balancing the Extremes of Educator Responses

and creatively and respond with flexibility, agility, and innovation within the context of an uncertain and complex world. These ends are our lodestar. Certain applications of generative AI will work directly against this mission. By pursuing universal adoption, and conflating substitutes for reasoning with gains to efficiency, uncritical advocates ignore real costs to their students—to say nothing of potential costs identified in other important debates, like increased misinformation and polarization or undermining confidence in scientific knowledge.¹¹

Recommendations for Educators

These two diametrically opposed perspectives about generative AI are likely to distort educators' responses in the classroom. Structuring the debate as volleys between such contradictory positions will foster confusion among educators hoping to navigate this new complexity while still mastering their defined tasks.

Educators will benefit from acknowledging both the important benefits of AI and the need for guardrails on its use, two imperatives clarified by attending to the mission of PME educators and to the specific applications of generative AI. Informed by this perspective, the following set of practical recommendations will help guide educators as they engage with AI to capitalize on gains to productivity while supporting their goal of developing critical thinkers.

Recommendation: Seize the Opportunity to Reevaluate and Revamp Pedagogy and Assessment to Engage Students

This new environment presents an opportunity to reevaluate pedagogy and assessments. Instead of working to “catch” student use of AI, educators should work to rethink instruction and assignments to inspire creativity, activate passion for course material, and incentivize engagement with the learning process. The key is to think in terms of fostering student engagement with their learning and assignments, such that students would prefer not to offload all work onto a chatbot, rather than in terms of revealing a chatbot's contributions.

Educators should reevaluate their written assignments seriously. Assignments that can be accomplished “effortlessly” by AI without meaningful human engagement should be modified or eliminated. For example, a chatbot with little student engagement can complete an assignment that asks for a summary of a famous historical event. Meanwhile, an assignment that asks students to select a historical event that is important to their family or own military experiences and interpret it using course concepts will be doable for a chatbot but will discourage a student from making that choice. A natural reaction here is displacement (“it is my students' responsibility not to use AI”). Now more than

ever, it may fall to students to find the motivation to own their learning experience. But educators can still play a role by setting demanding expectations and challenging their students to grow and overcome discomfort. A useful, albeit demanding, starting assumption for instructors might be that every existing assignment can be made less formulaic and more tailored to activate students' core motivations, thus disincentivizing counterproductive uses of AI and incentivizing the personal pursuit of knowledge.

Educators should take a similar approach in terms of pedagogy. Are there ways to conduct classroom instruction to encourage good use of generative AI? For example, can generative AI be used as a brainstorming function prior to deeper critical analysis in the classroom? The introduction of generative AI presents an opportunity for both educators and students. Educators who understand both the positives and negatives of the technology can revamp curriculum and ways of teaching to encourage greater student engagement.

Recommendation: Require All Course Syllabi to Address Appropriate Use of Generative AI in a Nuanced Manner

All course developers should be responsible for adding a generative AI provision in course syllabi to establish guardrails for its use on assignments and in the classroom. These provisions must be thoughtful. Provisions within a syllabus should not focus solely on limiting the use of ChatGPT (or specifying rules to allow for detection) but should also focus on structuring rules to actively promote learning and thoughtful use of AI.

A way to approach these provisions is to identify specific cases where AI use may be acceptable and where it may not. It is unrealistic that any set of provisions will capture all benefits or mitigate all risks, especially as the technology continues to advance; however, instructors who choose not to set any guardrails will create far greater risks to their students' learning outcomes and thus to PME and the Army as a whole.

A statement prohibiting certain uses of AI might look like this: "The use of generative AI is likely to inhibit the achievement of course objectives by undermining research and writing processes, which are

essential to the development of students' higher-order thinking skills. Therefore, the use of ChatGPT or other generative AI tools at any point during the research and writing process for the F100 Force Management paper is prohibited."

Statements promoting specific uses of generative AI might look like this: "Use of ChatGPT and other generative AI tools is acceptable and encouraged for brainstorming topics for this course's creative writing assignment, for use as a conversational tutor on core concepts from the course, and as coding aids for conducting quantitative analysis. Cite any contributions of generative AI tools. Keep in mind that these tools may provide misleading and incorrect responses, and ensure you avoid overreliance on AI output."

Recommendation: Avoid Over-testing with AI Content Detection Tools. Prioritize Instructors' Expert Intuition

Even when educators decide to prohibit AI use in specific cases, they must be careful to avoid excessively prioritizing detection of AI-generated content. Attending to possible academic violations is important, but overemphasizing detection entails a punitive and transactional mindset inconsistent with the core mission of educators. Moreover, testing randomly or wholesale for the use of AI in submitted work will cause educators to grossly overestimate both the prevalence of AI use in their classes and the likelihood of AI use in any specific piece of student work (see the sidebar on page 6).

Recommendation: Design, Implement, and Sustain Education and Training for Instructors on Generative AI

To answer questions of how AI use affects the development of critical-thinking skills and whether faculty should prohibit it or promote it in certain situations, it is essential for educators to understand the technology itself—what it is, what it can do, what it cannot do, and how it might evolve.

All faculty should have access to education and training on generative AI to build this knowledge. The term "training" here is key. Training on AI is different from a seminar-style faculty workshop that only discusses the

Analysis: AI Content Detection Tools' Susceptibility to False Positive Classifications of SS307 Papers

In April 2023, the authors analyzed the performance of three popular, publicly available AI content detection tools—Open AI's "AI Classifier," "GPTZero," and Copyleaks' "AI Content Detector"—against fifty student-written papers from a fall 2022 writing assignment for a political science course at the U.S. Military Academy and thirty ChatGPT responses to the assignment prompt.

The detection tools varied substantially in their performance. AI Classifier performed most conservatively, only detecting AI-generated text for 20 percent of ChatGPT outputs and never falsely flagging student text as AI-generated. (These are estimates from a small sample; OpenAI reported a 9 percent false positive rate from its own evaluations.¹) Meanwhile, we measured the Copyleaks tool's false positive rate at 4 percent. GPTZero was especially prone to false positives. GPTZero incorrectly labeled twenty-one of fifty student-written texts as possibly or likely AI-generated, yielding a false positive rate of 42 percent.

Classification Performance (True Positive Rates and False Positive Rates)

AI Classifier					Copyleaks					GPTZero							
		Predicted				Predicted				Predicted							
		Positive	Negative			Positive	Negative			Positive	Negative						
Actual	Positive	6	24	TPR:	0.200	Actual	Positive	26	4	TPR:	0.867	Actual	Positive	30	0	TPR:	1.000
	Negative	0	50	FPR:	0.000		Negative	2	48	FPR:	0.040		Negative	21	29	FPR:	0.420

(Figure by authors)

Instructors unaware of these misclassification rates will be prone to overestimating the prevalence of AI use. An instructor who randomly tests student work against GPTZero and receives the result that the text "may include" or "is likely entirely" AI-generated content may conclude that the student certainly used ChatGPT. However, given our measured classification rates, this would be wrong.

Moreover, even when instructors are aware of this potential for false positives, they will be prone to excessive suspicion of student work if they test papers randomly or wholesale against available detection tools, because we often neglect to account for base rates. An instructor who is aware of our measured 42 percent false positive rate for GPTZero might conclude that a flagged random sample is AI output with 58 percent certainty. However, given our measured classification rates, and when accounting for an assumed base rate like 2 percent student use of AI on the assignment, the actual likelihood of that writing sample containing AI content is closer to 5 percent.

Note

1. "New AI Classifier for Indicating AI-Written Text," OpenAI, 31 January 2023, <https://openai.com/index/new-ai-classifier-for-indicating-ai-written-text/>.

technology in abstract terms. Training requires prepared instruction on how the technology works and what it does. For example, West Point hosted training on large language models and their capabilities and constraints in January 2023 to help faculty determine when it would be appropriate to allow the use of ChatGPT in the classroom.

These efforts at West Point are commendable, but if not sustained, they will be incomplete. All PME's should consider instituting periodic training and conversations on ChatGPT and other AI developments over the long run. The institutional response cannot be a short-term frenzy followed by long-term silence. Just as generative AI models change whenever they encounter new "realities" in their training data, faculty members should regularly update their understanding of AI tools.

Recommendation: Increase Institutional Messaging at All Levels of Academic Leadership about the Core Mission in the ChatGPT Era

All faculty at PME's bear a burden of addressing the use of generative AI in the classroom and on assignments. Bottom-up solutions are essential in this domain. However, faculty may not fully understand the technology or its costs and benefits, and they may struggle to advocate even for well-founded recommendations in the face of loud criticism from others occupying the extremes.

Consistent messaging, in written policy or verbal communication, from PME leaders will support faculty attempting to navigate the sea change. Regardless of any new technology (or pandemic or world crisis), the purpose of Army PME remains constant—to educate agile and adaptive leaders who can achieve cognitive overmatch to win in complex environments.¹² Discussions around AI should continually raise questions addressing how this specific use of this technology will promote officers'

learning and development of higher-order thinking skills. Generative AI will be beneficial for education, but not in every circumstance, and faculty will be better equipped to identify these specific circumstances when they orient on the purpose of professional military education.

What the Future Can Tell Us about Our Current Approach

Things are moving fast. Within the past year, Google announced the initial release of Bard, now Gemini, its "experimental conversational AI service," and observed that "the scale of the largest AI computations is doubling every six months, far outpacing Moore's Law."¹³ Steve Wozniak, Elon Musk, and other tech leaders called for a pause on the "out-of-control race" in AI development.¹⁴

Given such a high rate of change, it seems plausible that even the most thoughtful, personalized, complex, or contemporaneous assignments will be doable for chatbots in the not-too-distant future. That future, and the contentiousness of debates engendered by visions of it, make it tempting to throw up our hands. However, with the Department of Defense's Data Strategy encouraging greater integration of AI, educators in PME should embrace the opportunities to lead in this space. Educators who commit to generating endlessly more convoluted assignments just to discourage AI use will commit themselves to a Sisyphean fate. On the other hand, it will also remain counterproductive to be uncritically accepting of AI as a "golden touch" across all cases. Neither fearful nor ignorant of the technology, PME educators should strive to identify valuable use cases and revamp their pedagogy, preserve trust with their students, and remain confident in their mission of developing the critical-thinking abilities of the Army's leaders. ■

Notes

1. Yusuf Mehdi, "Reinventing Search with a New AI-Powered Microsoft Bing and Edge, Your Copilot for the Web," Official Microsoft Blog, 7 February 2023, <https://blogs.microsoft.com/blog/2023/02/07/reinventing-search-with-a-new-ai-powered-microsoft-bing-and-edge-your-copilot-for-the-web/>; Josh Achain et al., "Open AI (2023): GPT-4 Technical Report," OpenAI, 15 March 2023, <https://cdn.openai.com/papers/gpt-4.pdf>.

2. Stephen Marche, "The College Essay Is Dead," *The Atlantic* (website), 6 December 2022, <https://www.theatlantic.com/technology/archive/2022/12/chatgpt-ai-writing-college-student-essays/672371/>; Jason Wingard, "ChatGPT: A Threat to Higher Education?," *Forbes* (website), 10 January 2023, <https://www.forbes.com/sites/jasonwingard/2023/01/10/chatgpt-a-threat-to-higher-education/?sh=7711e8b31e76>.

3. Michael Elsen-Rooney, "NYC Education Department Blocks ChatGPT on School Devices, Networks," Chalkbeat, 3 January 2023, <https://www.chalkbeat.org/newyork/2023/1/3/23537987/nyc-schools-ban-chatgpt-writing-artificial-intelligence>.

4. Emma Bowman, "A College Student Created an App That Can Tell Whether AI Wrote an Essay," NPR, 9 January 2023, <https://www.npr.org/2023/01/09/1147549845/gptzero-ai-chatgpt-edward-tian-plagiarism>.

5. Ezra Klein, "The Surprising Thing AI Engineers Will Tell You If You Let Them," *New York Times* (website), 16 April 2023, <https://www.nytimes.com/2023/04/16/opinion/this-is-too-important-to-leave-to-microsoft-google-and-facebook.html>.

6. Jackie D. Kem, "Command and General Staff College Academics Ethics Policy," CGSC Bulletin 920, 3 August 2023.

7. U.S. Department of Defense (DOD), *DoD Data Strategy* (Washington, DC: U.S. DOD, 30 September 2020), <https://media.defense.gov/2020/Oct/08/2002514180/-1/-1/0/DOD-DATA-STRATEGY.PDF>.

8. David G. Savage, "Calculator Issue: Math Class; Old Ways Don't Add Up," *Los Angeles Times* (website), 17 May 1986, <https://www.latimes.com/archives/la-xpm-1986-05-17-mn-5638-story.html>; Anndee Hochman, "Math Teachers Stage a Calculated Protest," *Washington Post* (website), 4 April 1986, <https://www.washingtonpost.com/archive/local/1986/04/04/math-teachers-stage-a-calculated-protest/c003ddaf-b86f-4f2b-92ca-08533f3a5896/>; Todd Oppenheimer, "The Computer Delusion,"

The Atlantic, July 1997, <https://www.theatlantic.com/magazine/archive/1997/07/the-computer-delusion/376899/>.

9. Andrej Karpathy (@karpathy), "The hottest new programming language is English," X (formerly Twitter), 24 January 2023, 2:14 p.m., <https://twitter.com/karpathy/status/1617979122625712128?lang=en>.

10. Ido Roll and Ruth Wylie, "Evolution and Revolution in Artificial Intelligence in Education," *International Journal of Artificial Intelligence in Education* 26 (2016): 582–99, <https://doi.org/10.1007/s40593-016-0110-3>.

11. Leah Thomas, "New Report Finds Bias in ChatGPT," Manhattan Institute, 14 March 2023, <https://manhattan.institute/article/new-report-finds-bias-in-chatgpt>; Blayne Haggart, "Unlike with Academics and Reporters, You Can't Check When ChatGPT's Telling the Truth," *The Conversation*, 30 January 2023, <https://the-conversation.com/unlike-with-academics-and-reporters-you-cant-check-when-chatgpts-telling-the-truth-198463>.

12. "About ArmyU," Army University, accessed 30 April 2024, <https://armyuniversity.edu/About>.

13. Sundar Pichai, "An Important Next Step on Our AI Journey," *The Keyword* (blog), Google, 6 February 2023, <https://blog.google/technology/ai/bard-google-ai-search-updates/>.

14. "Pause Giant AI Experiments: An Open Letter," Future of Life Institute, 22 March 2023, <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>.