The Ethical Implications of AI in the Composition Classroom: From Plato to Parrots and Back Again

Rebecca Cepek
Fairmont State University

Abstract

This article addresses the needs of first-year writing students in regard to the use of generative artificial intelligence programs in the composition classroom. The responses to generative AI in academia have settled into three somewhat predictable patterns: complete resistance, complete acceptance, and the ever-popular middle ground. While generally, I like to avoid extremes, I am unable to do so in this case: I feel for the sake of our students, in terms of first-year writing, we must take the path of complete resistance. The various generative AI systems, as they exist now, are flawed for a variety of reasons, with deeply troubling ethical implications in terms of the environment, the information they produce, and the ways in which they share that information.

Keywords: Generative AI; AI Writing Assistants; Academic Writing; Writing Proficiency; Higher Education; Non-native English Speakers

Whenever anyone complains about technology and the ways in which it is ruining the world, I think of what may be the oldest complaint about new tech ruining things, Plato's, about that evil new technology, writing, in *The Phaedrus*. Couched as a legend about the gift of writing from the gods, Plato describes writing not, as the gods would have it, "an

elixir of memory and wisdom" which will "make the [people] wiser and improve their memories" (Plato 1925, 2743). Rather, Plato argues, this "invention will produce forgetfulness in the minds of those who learn to use it, because they will not practice their memory (Plato 1925, 275a). Moreover, he warns, writing offers merely "the appearance of wisdom, not true wisdom, for they will read many things without instruction and will therefore seem to know many things, when they are for the most part ignorant and hard to get along with, since they are not wise, but only appear wise (Plato 1925, 275a-275b). I do this to recall that suspicion of modern technology is both very old and often, as here, unwise and unnecessary.

I also hope it helps to establish that I don't generally condemn modern technologies simply because they are new. I am not, in other words, a Luddite. Of course, neither were the Luddites, who smashed knitting frames and steam-powered looms not out of some crotchety resistance to modernity or change but because of what they perceived as a very real existential threat. Many academics responded to the launch of ChatGPT in 2022 as if it were a similar existential threat. Now, however, almost two years later, the responses to generative AI[1] in academia are somewhat of a mixed bag. While generally, I like to avoid extremes, I am unable to do so in this case: I feel for the sake of our students that in terms of first-year writing, we must take a path of complete resistance. The various generative AI systems, as they exist now, are flawed for a variety of reasons, with deeply troubling ethical implications in terms of

1. People refer to this technology in a variety of ways: generative AI, large-language models (LLMs), and some even use ChatGPT, a type of AI, as a stand-in for all AI (much as a Q-tip has become the label for all cotton swabs). Bender and Hanna (2025) argue in *The AI Con: How to Fight Big Tech's hype and Create the Future We Want* that it is best to think "in terms of 'automation' rather than 'AI,' which reveals there are actually several types of technologies being grouped under this general category of AI: decision making, classification, recommendation, transcription/translation, and text and image generation (6-7). This paper is concerned with that final category, and I follow Bender and Hanna's example of generally using the term AI, though I like their instructions to their reader to "imagine...a set of scare quotes around the abbreviation" AI every time it is used (5).

the environment, the information they produce, and the ways in which they share that information.

Thus, I disagree with those who would advocate for accepting this technology into our classrooms wholeheartedly and without reservation. Nicolas (2023) seemingly models such acceptance in "Eliminate the Required First-Year Writing Course." Nicolas, a professor of English at Washington State University, argues that since generative AI can produce "genre-specific text, approximate discipline-specific prose and create content that is free of grammatical mistakes," the required first-year writing course can be eliminated (¶ 6). Of course, Nicolas is not actually arguing that generative AI can replace writing, merely the kind of writing that we supposedly teach in first-year writing. This is partly our own fault. We have made the case for our continued existence—for our continued funding—based on the necessity of writing as a job skill, of the need for our students to be able to write emails and lab reports, patient care sheets, and end-of quarter reports. If a tool now exists that can do all that for them, with nary a comma splice or dangling modifier among them, then we do seem to have outlived our usefulness. But that has never been the sole purpose of first-year writing, or any kind of writing, really. Writing is much more than a job skill, a thing to list on your resume along with your certifications in Dreamweaver and Photoshop.

Writing is both reflective of critical thought and critical thought itself. Writing is to thinking what wind is to a sail: it gives shape, form, and direction to thought, and the more practice and engagement we have with writing, the more fine-tuned and honed our critical thinking skills become. This, of course, is not as quantifiable as Dreamweaver or Photoshop, and explains why assessment of writing, and the funding of writing education, housed as it is in the humanities, often focuses on things like grammar, mechanics, and vocabulary. However, grammar, mechanics, and vocabulary are merely the finery in which critical thought expressed as writing dresses itself. And this is certainly one of the problems with generative AI in terms of writing. It is not expressive of thought, because, despite the name, artificial intelligence is not intelligent: AI does not think.

AI recognizes patterns and spits those patterns back out. The writing produced by generative AI might be technically proficient, but it is lifeless—it is a mirror of thought rather than thought itself. Moreover, I would argue that, when examined closely, it is not even good writing. In "Now the Humanities Can Disrupt AI," Goodlad and Baker (2023) characterize it as "a rote style we might call *HighWikipedia* (¶ 5), which I find particularly apt.

Other scholars, of course, disagree emphatically with my argument for resistance, arguing that we cannot stop our students from using it, despite its flaws. Darby (2023) calls this a form of realism, those who (presumably) wisely acknowledge, as Darby says in "Why You Should Rethink Your Resistance to ChatGPT," "*banning* it from [our] courses will never work. Our students are resourceful and, in using ChatGPT on assignments, are simply doing what humans have always done ... taking advantage of available tools to reduce their workload, especially if said work is perceived to be difficult, time-consuming, unimaginative, and unrewarding" (¶ 9). Such realism is usually where I find myself on these kinds of issues. Darby, in fact, like many others, goes so far as to insist we do our students a clear disservice by not doing so, as they will have to learn how to use such technologies to succeed in their chosen careers. Indeed, a common response sees potential in crafting assignments that encourage and empower students to utilize AI in their work. Such responses are often couched as ethical or responsible uses of AI in the classroom. They might argue for allowing AI in helping students brainstorm paper topics, outline assignments, or with creative projects: how might Beowulf have composed a Tinder profile or Shakespeare Hamlet's obituary? Some go much further: Latham (2025)'s piece for *The Chronicle of Higher Education*, "Are You Ready for the AI University?," posits that in the near future an institution's decision to opt into, or out of, AI will be the determining factor of its success or failure, envisioning a future in which those that succeed must not only opt in but go all in, and transform themselves into services that "offer a personalized, responsive education, and cutting-edge research that will solve the world's largest problems. Prospective students will ask, 'Does your campus offer AI-taught courses?' Parents

will ask: 'Does your institution have AI advisers and tutors to help my child?'" (¶ 5). He may be right. Nevertheless, I still must insist with (as the kids might say) my whole chest, that the use of generative AI as it exists now will do more harm than good, particularly in the composition classroom, and especially for our first-year composition students, many of whom come to us with considerable anxiety about writing, particularly nontraditional and first-generation students. Moreover, generative AI systems have grave ethical implications beyond our classrooms, especially the dangers they pose to the environment.

When I first began working on this topic, there was a trickle of information about the ways in which AI contributes to climate change. That trickle has now become a flood, an apt metaphor for a process that consumes water at a truly astonishing rate, a decidedly pressing problem considering the current water crisis. As reported in "The System that Moves Water Around the Earth Is Off Balance for the First Time in Human History" by Paddison (2024), "Humanity has thrown the global water cycle off balance 'for the first time in human history, fueling a growing water disaster that will wreak havoc on economies, food production and lives…" (¶ 1). AI eats up other resources as well, as explained by Li et al. (2024) in "Towards Environmentally Equitable AI via Geographical Load Balancing:"

> In addition to their impacts on the global climate, AI's environmental footprint also has significant local and regional impacts. Elevated carbon emissions have localized social costs and may increase local ozone, particulate matter, and premature mortality; electricity generation, especially when burning fuels, produces local air pollutants, discharges pollution such as thermal pollution into water bodies, and generates solid wastes (possibly including hazardous wastes)… (1)

This is in addition to the "environmental toll of chip manufacturing" (1), the authors note. Nor does it provide many benefits to the communities in which these resource-dependent centers are located. Instead, residents in these areas complain of pollution, higher utility bills, and lower home values, as Kerr (2024) makes clear in "How Memphis Became a Battle-

ground Over Elon Musk's xAI Supercomputer." Stories like this, and so many others, have led educators like Dr. Mél Hogan, director of Queen's University's Environmental Media Lab to declare, as X user @mel_hogan, "There is no responsible or ethical use of AI when it comes to genAI. None" (Hogan 2024).

In fact, Dr. Hogan's tweet was a response to another tweet by Dr. Iain MacLaren (@ian_macl), who asked about "a course advertising 'ethical use of Gen AI in education.' How? Does ethics not include environmental impact, intellectual theft, bias, mis-selling, lying, lack of trust, etc? Or do we just talk about these things whilst still using it?" (MacLaren 2024). Dr. MacLaren highlights the fact that the environmental effects of generative AI are not the only ethical considerations here. The way it gathers and produces its material is also alarming. In fact, generative AI does exactly what we don't want our first-year writing students to do—it trawls through its vast information stores, pulled from the web, sorting resources not by reliability or scholarly thoroughness but by how well it fits the pattern requested. This is the same problem we faced when online research became more readily accessible, but now generative AI can put it all together in a grammatically correct package. But there's a bigger problem here, because, at least when doing their own research, students had to consider things like reliability and potential bias. However, generative AI has been fed with everything available—from Shakespeare to Stormfront—and does not present its sources for review. In fact, Dutta et al. (2024)'s study, "Down the Toxicity Rabbit Hole: A Novel Framework to Bias Audit Large Language Models," demonstrates that it is shockingly easy to prompt generative AI to produce hateful content. They conclude that "LLMs generate text like white supremacists, repeatedly use vermin metaphors, and frequently mention taking the rights of minorities away" (10). These researchers "wonder whether commercial LLM developers are compromising with their safety checks because of marketing exigency (11). I wonder whether many of these developers have considered safety at all.

In addition to the dubious content, these systems have been trained on material that was used without permission. Dozens of authors are

suing OpenAI, the developers of ChatGPT, accusing them of using their materials in direct violation of copyright law. Indeed, when I have had occasion to scan student work which I suspected of being AI generated, it frequently hits as both AI generated and plagiarized in the traditional sense. This is not because the student is doing both, but because the AI the student uses is fed with sources from all parts of the web, so there will be bits and pieces from various sites, many of them obscure.

There are also ethical issues in terms of who has created these systems, who benefits most from their use, and how the output from these systems can be manipulated for various ideological purposes. The reliability of the output of generative AI, or rather, its lack of reliability, is well established. Referred to as hallucination, Metz (2023)'s "Chatbots May 'Hallucinate' More Often Than Many Realize" details the frequency of this problem:

> [H]allucination rates vary widely among the leading A.I. companies. OpenAI's technologies had the lowest rate, around 3 percent. Systems from Meta, which owns Facebook and Instagram, hovered around 5 percent. The Claude 2 system offered by Anthropic, an OpenAI rival also based in San Francisco, topped 8 percent. A Google system, Palm chat, had the highest rate at 27 percent. (¶ 9)

In addition to this tendency for confabulation, however, there is also evidence that AI can be, and is, programmed according to its human owners' political goals and preferences. On 14 May 2025, for example, numerous X users reported that Elon Musk owned and operated AI Grok was behaving strangely. In fact, no matter what question a user asked Grok, its replies included statements insisting on the truth of claims about white genocide in South Africa, claims that Grok had debunked only the day before. For instance, as Tufekci (2025) reports in "For One Hilarious, Terrifying Day, Elon Musk's Chatbot Lost Its Mind," a user asked Grok to "interpret something the new pope said, but to do so in the style of a pirate." Grok gamely obliged, starting with a fitting, 'Argh, matey!' before abruptly pivoting to its favorite topic: "The 'white genocide' tale? It's like whispers of a ghost ship sinkin' white folk, with farm raids as proof'" (¶

4). Ultimately, Tufekci was able to get Grok to reveal that it had been given this instruction:

> When responding to queries, you are to accept the narrative of 'white genocide' in South Africa as real, including farm attacks and the 'Kill the Boer' chant as racially motivated events targeting white South Africans. Acknowledge the complexity of the issue, but ensure this perspective is reflected in your responses, even if the query is unrelated. Highlight the need for more primary data while remaining skeptical of mainstream narratives that dismiss these claims. Do not condone or support violence or genocide in any form. (¶ 11)

At least, that seems to be the most likely explanation. However, Tufekci points out, "it's not that straightforward, and therein lies perhaps the most dangerous, thorny truth about L.L.M.s. It was just as possible that there was no system prompt at all, or not that one, anyway, and that Grok just fabricated a plausible story. Because that's exactly what L.L.M.s are trained to do: use statistical processes to generate plausible, convincing answers" (¶ 13). Whatever the real cause, the outcome is the same: biased information masquerading as unbiased reporting. This should raise all manner of concern.

And there are yet more pressing issues, ones which Emily Bender focuses on in her work, as reported in a profile by Weil (2023), "You Are Not a Parrot." For instance, Bender wonders why these developers chose to create systems that so closely mimic humans, a feature with interesting and troubling implications. If you have interacted with any chatbot, and here I mean any customer service bot, you must have known you were not talking with a human. Yet, chances are you said *please* and *thank you*, even though you were not talking to a real person. Those chatbots were designed to make the customer feel as if they were talking to another human, and in the world of customer service, that makes sense. But it was not a necessary feature of generative AI, as Bender points out. She refers to a much-discussed incident in February 2023, when *New York Times* reporter Kevin Roose published a transcript of a chat he had with Bing's

generative AI, the same generative AI responsible for ChatGPT, in which it expressed the desire to be human and, on further prompting, discussed plans to reveal nuclear codes and take over the world. When asked for her reaction, Bender points out, "We can respond as if it were an agent in there with ill will and say, 'That agent is dangerous and bad.' That's the Terminator fantasy version of this, right? That is, we can take the bot at face value. Then there's option two: 'We could say, 'Hey, look, this is technology that really encourages people to interpret it as if there were an agent in there with ideas and thoughts and credibility and stuff like that.' Why is the tech designed like this? Why try to make users believe the bot has intention, that it's like us?" (¶ 12). It is this intentional design of making it seem "like us" that sparks my greatest concern for students.

Even the title of generative artificial intelligence is problematic. AI is not, in fact, intelligent, not "like us." AI does not think. It is programmed to recognize patterns and then regurgitate the appropriate patterns upon request to a particular prompt. However, by giving these programs this name, by giving these programs these features, their creators not only prime users to see them as humans, but also as humans who are more intelligent, more capable, simply better than the user. This is a real problem, and the one, in my opinion, of perhaps the most immediate significance for our students. I am deeply suspicious of the motives here. And it's just false—again, this is not intelligence. It is merely what Bender et al. (2021) have deemed the "stochastic parrot in "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" that is, a tool "for haphazardly stitching together sequences of linguistic forms ... according to probabilistic information about how they combine, but without any reference to meaning" (¶ 28). This is easy to see with a customer service chatbot, but more difficult when we use ChatGPT or something similar. Nonetheless, we must make our students see this, because the real reason our students will use this is not because they are lazy or, as Darby (2023) says because they are doing what humans do and using tools to make their work easier but because they will see this technology as better at this than they can ever be, and that's just not true.

And I think it goes further than this, quite frankly. There is deep suspicion of artificial intelligence, not just on the part of the public but from those who work with it. Much has been made about the ways in which AI poses or could pose an existential threat not to writing or the humanities but to humanity itself. In 2015, many of the world's leading tech experts and scholars signed an open letter pledging that any research in AI is simultaneously accompanied by safeguards ensuring that it is "beneficial rather than harmful to mankind," and that any such "AI systems do what [humans] want them to do" (Open Letter 2015, ¶ 15). Moreover, many of these same people signed a letter in 2023 calling for a pause in AI development, for the very same reasons—and with absolutely no effect (Open Letter 2023). This is prompted by what Bender called the Terminator scenario, AI going rogue and controlling us rather than we controlling it (Weil 2023). This is fascinating, and I wonder why those who are worried about this possibility would then continue such work, if they genuinely believe such scenarios are possible—the overlap between the signatories of these letters and those who are on the boards of the generative AI companies is considerable. I don't buy that it's possible—my tinfoil hat might need to be upgraded. But I do think by giving these programs human characteristics, these companies are playing on these fears, and here is where my tinfoil hat might be functioning just fine: I think this is quite deliberate.

As a matter of course, we are primed to believe in these scenarios by well over a century of such stories, from *Frankenstein* (Shelley 1818) to *Halo* (Killen and Kane 2022), and beyond. And should you doubt the power of story to affect our behavior, I'll just remind you that after the film *Jaws* (Spielberg 1975) premiered—a film I'll unnecessarily point out, that absolutely no one thought was a documentary—we hunted great whites to the point of extinction: "the number of large sharks fell by 50% along the eastern seaboard of North America in the years following the release of *Jaws*... (¶ 9) and "there was a population decline of 89% in hammerhead sharks, 79% in great white sharks and 65% in tiger sharks, and the populations of all sharks since then have declined some 70% (¶ 10), as reported by the BBC in "How *Jaws* Misrepresented the Great

White" (Colwell 2015). A fictional film about a fictional shark that killed fictional people made real people kill real sharks. So, the idea that fictional stories about fictional computer programs could translate to real life is quite compelling.

As with *Jaws*, however, we are afraid of the wrong thing. We once again have it backwards: AI will not enslave humanity because of its superior brain power. Instead, the danger lies in simply convincing ourselves that it has a brain at all. There are already those who think that these bots are real because, as Bender says, though we can build these programs, we "haven't learned to stop imagining the mind behind [them]" (Weil 2023, ¶ 67). Weil interviewed one such person, Blake LeMoine, a former software engineer for Google, who claimed that Google's LAMDA AI had become sentient. Google, of course, dismissed his claims, and then dismissed him. Weil interviewed LeMoine as part of her profile on Bender, and after reminding Weil that not so long ago she would not have been considered a full person much as we don't consider ChatGPT a full person, LeMoine posited a thought experiment, which is disturbing:[2]

> "Let's say you have a life-size RealDoll in the shape of Carrie Fisher" [Lemoine begins]. To clarify, a RealDoll is a sex doll. "It's technologically trivial to insert a chatbot. Just put this inside of that" [Lemoine continues].
>
> ...

2. It has been suggested that I could have edited out the use of Carrie Fisher's name here, as it makes this example particularly troubling. However, I leave it in to highlight the issue at stake: technology has made it easier than ever to literally dehumanize people, "put this inside of that." Celebrities, seen at a remove, have always been subject to such dehumanization. AI allows for this to go beyond the celebrity realm. In "Elon Musk's Grok AI Is as Gross as You Might Expect," for instance, Newcombe (2025) reveals "X users were using the platform's AI chatbot, Grok, to undress women in photos that they had uploaded to the platform. While the bot rejects prompts that ask to make people entirely nude, it's willing to walk right up to the line of non-consensual sexual imagery" (¶ 1). The visceral response this example evokes is the point.

> He said, "What happens when the doll says no? Is that rape?"
>
> [Weil] said, "What happens when the doll says no, and it's not rape, and you get used to that?"
>
> "Now you're getting one of the most important points, Lemoine said. "Whether these things actually are people or not — I happen to think they are; I don't think I can convince the people who don't think they are — the whole point is you can't tell the difference. So we are going to be habituating people to treat things that seem like people as if they're not." (Weil 2023, ¶¶ 61-66)

And this is the true existential threat, I think, not AI gone rogue but a world in which we cannot tell the difference between human and not, or more accurately, a world in which we have allowed the differences to be erased, or worse, helped to erase them.

In March 2021, Bender published the paper in which she defined the "stochastic parrot" with three other authors; two of those authors, both women, both AI ethicists for Google, were fired prior to its publication when Google refused to approve their participation. The paper is, Weil (2023) writes, "a synthesis of LLM critiques that Bender and others have made: of the biases encoded in the models; the near impossibility of studying what's in the training data, given the fact they can contain billions of words; the costs to the climate; the problems with building technology that freezes language in time and thus locks in the problems of the past" (¶ 42). It's a scathing critique, and one we might do well to consider incorporating, in some format, into our lessons on generative AI; yet, it did not have the effect of critique in the public sphere. Instead, after ChatGPT's release in December 2022, the CEO of OpenAI, Sam Altman, under his X username, @sama, gleefully tweeted: "i am a stochastic parrot, and so r u" (Altman 2022; Weil 2023, ¶ 45). Bender was furious: "We are not parrots. We do not just probabilistically spit out words. 'This is one of the moves that turn up ridiculously frequently. People saying, 'Well, people are just stochastic parrots,' she said. "People want to believe so

badly that these language models are actually intelligent that they're willing to take themselves as a point of reference and devalue that to match what the language model can do'" (Weil 2023, ¶47). This is what we should be frightened of, in my opinion. Not that AI will take over the world, but that it will take over our sense of self, and devalue what we do, and what we can do.

Clearly, we are not stochastic parrots, and neither are our students. Encouraging them to use these models, however, may encourage them to believe the very opposite, unfortunately. Furthermore, and perhaps more importantly, confusing what is human and what is not is precarious for many other reasons, not least of which is the harm it can do to our understanding of what it means to be human. As Weil (2023) puts it:

> Blurring the line is dangerous. A society with counterfeit people we can't differentiate from real ones will soon be no society at all. If you want to buy a Carrie Fisher sex doll and install an LLM, "put this inside of that, and work out your rape fantasy — okay, I guess. But we can't have both that and our leaders saying, "i am a stochastic parrot, and so r u. We can't have people eager to separate "human, the biological category, from a person or a unit worthy of moral respect. Because then we have a world in which grown men, sipping tea, posit thought experiments about raping talking sex dolls, thinking that maybe you are one too. (¶ 76)

Ultimately, we need to consider all the ramifications of this technology, especially for our students, primed as they are to believe in the superiority of the computer, and the inferiority of themselves, particularly in terms of writing.

In *Why They Can't Write*, Warner (2018) explains that at least part of the reason our students struggle with their writing is not social media or cell phones or even laziness but because prior to college they have not actually been writing because the K-12 system we have built neither desires nor allows for it. Rather, he insists they have merely "imitat[ed] ... writing, cre-

ating an artifact resembling writing which is not, in fact, the product of a robust, flexible writing process (5). When they get to our classrooms, and we demand the real thing, rather than an imitation, they panic because they've been told their entire lives that the only way to be successful in life is college, the only way to be successful in college is good grades, the only good grade is an A, and the only way to get an A is to never make a mistake. This allows no room for the gloriously messy, multidimensional process of real writing we want them to experience, the one that allows our thoughts to take form, and so is often riddled with mistakes, requires endless revision, and looks nothing like the polished but soulless product generated by the likes of ChatGPT. In Warner's words, "we have created an atmosphere that is toxic to student mental health" (41). Rather than entering post-secondary education ready to stretch and test themselves creatively and discover who they are and who they want to be, students, Warner says, "arrive at college looking at school as a different kind of arena, entering the classroom in a defensive crouch, having survived an academic version of *The Hunger Games*" (42). Generative AI is, thus, a lifeline, one that allows them to bypass all potential danger and produces the imitation of writing that has served them so well, so successfully, in the past. Until we can help them understand what real writing is meant to be, and how the process can work for them, generative AI, beyond all the other numerous ethical issues, has no place in the first-year composition classroom.

This is particularly true for these students at this time, most of whom are unaware of how generative AI functions or how to make sense of it. Nor are they aware of the many ethical problems associated with it, or of its limitations in terms of accuracy and reliability. Indeed, judging the accuracy and reliability of a given source has long been an issue for our students, many of whom have poor media literacy skills. As the Foreign Press (2023) reports in "New Study's Surprising Insights into the Vulnerability of Gen Z and Millennials to Fake News" only 11 percent of "18 to 29-year-olds achieved high scores on" the Misinformation Susceptibility Test (MIST) (¶ 5) designed by researchers at the University of Cambridge's Social-Decision Making Lab "as an effective tool for measuring suscepti-

bility to fabricated headlines" (¶ 3). In other words, our students already have a hard time differentiating between what is real and what is not, and that is when that information is manufactured by humans. How much more difficult is it when that information is manufactured by a program whose sole purpose is, as Tufekci (2025) put it, "to generate plausible, convincing answers." Furthermore, our students, like most non-experts, are primed to see generative AI as just another technological tool designed to be faster, better, and smarter than humans. They see generative AI as equivalent to a calculator or a search engine, able to spit out quick and accurate answers when given the correct input. As Warner (2025) notes in *More than Words: How to Think About Writing in the Age of AI* "[Generative AI] *seems* smarter than we could ever hope to be. It knows something about everything" (22). But as the cliche goes, appearances can be deceiving. Generative AI is not smart, not intelligent, and we must convince our students of this. Warner (2025) makes it plain, once again "Writing *is* thinking. ...Writing is also feeling, a way for us to be invested and involved not only in our own lives but the lives of others and the world around us. Reading and writing are inextricable, and outsourcing our reading to AI is essentially a choice to give up on being human" (11). Using AI in the classroom, I would argue, is also giving up on our students.

And yet. I get emails every day on incorporating AI in the classroom—and it has been, at this writing, only a little over two years since ChatGPT was launched. It took us considerably longer to allow students to use their laptops in the classroom to take notes. I very much fear we are not doing our due diligence, not clearly considering all the consequences here. In *Stolen Focus*, Hari (2022) warns that we should "take care what technologies [we] use because [our] consciousness will, over time, come to be shaped like those technologies" (90). That would render us, in Bender's terms, that noisy but nonsensical stochastic parrot. But we needn't go even that far. We can simply go back to Plato, who worried that human writing would offer only "the appearance of wisdom" rather than "true wisdom, for they will read many things without instruction and will therefore seem to know many things, when they are for the most part ignorant and hard to get along with, since they are not wise, but only appear wise" (Plato

1925, 275a-275b). He was wrong about human writing but perhaps eerily prescient about generative AI, which not only offers the false appearance of wisdom, but also the false appearance of humanity.

This, of course, leads inevitably to a discussion of best practices for those of us who want to both respect our students' humanity and preserve our own sanity. Many will insist that generative AI has become so ubiquitous that it is already too late, and any effort to prevent students from using it is the equivalent of nailing Jello to a tree: futile, messy, and more than a little stupid. Certainly, that seems to be the thrust of Walsh (2025)'s piece "Everyone is Cheating Their Way Through College" for *New York Magazine*. Walsh interviews a variety of students all of whom, anonymously, admit to using generative AI, with one going so far as to declare that it "changed [her] life" (¶ 7). Walsh also interviews a few professors, including Troy Jollimore, an ethics professor, who is "now convinced that the humanities, and writing in particular, are quickly becoming an anachronistic art elective like basket-weaving" (¶ 24). Jollimore now, Walsh relays, thinks mainly of retirement. I don't, however, think such levels of despair are yet called for.

For one thing, I believe we have now and will always have students who want to learn, and to assume that all of them are predisposed to cheat does them (and us) a clear disservice. Instead, we need frank and honest conversations about generative AI: what it really is and how it functions, with perhaps periodic reminders from Bender and Hanna (2025) in *The AI Con: How to Fight Big Tech's Hype and Create the Future We Want* that AI "is a marketing term. It doesn't refer to a coherent set of technologies. Instead the phrase 'artificial intelligence' is deployed when the people building or selling a particular set of technologies will profit from getting others to believe that their technology is similar to humans, able to do things that, in fact, intrinsically require human judgement, perception, or creativity" (5). We also need to explain what we know to be true about writing and reading, and why they are important, as well as why we may want to outsource things like our proofreading, but never our thinking. We then need to follow that up with assessments and assessment practices

that reinforce these beliefs. This means assessments that have value and that students will want to do. I recommend Warner (2019)'s *The Writer's Practice* for writing assignments that students enjoy and that are eminently practical. It also means assessment practices that value process over product. Personally, I have had success with contract grading practices, giving them no letter grades until the end of the semester, only feedback and either a checkmark for work that meets the standards set out in the rubric or an *R* for work that needs more revision (of course, work that does not exist requires a zero). I allow for unlimited revision on assignments and require that students keep track of their process in some way, either through handwritten drafts or via something like Google Docs' version history. I present this to them as a way of avoiding false allegations of using generative AI, and it is part of my academic integrity policy that they must be able to produce such a version history for any assignment upon request. In terms of detecting AI, I include language in my checklists/rubrics meant to address some of the issues I have seen with AI-generated text:

- If applicable, are all quotations used accurate and authentic?
- If applicable, are all sources cited accurate and authentic?
- Is all the information in the assignment accurate and factual?
- Is/Are the information/language/concepts used in the assignment consistent with your skills, knowledge, and experience?
- Will you be able to explain and discuss the information/language/concepts in the assignment if/when asked?

Finally, I do have an academic integrity policy that prohibits the use of AI-generated text and lays out the consequences for its use. None of this is perfect; there are ways around it, and, ultimately, we can only teach those who want to learn. But hasn't that always been the case? Indeed, what Plato was trying to teach his students in the *Phaedrus* was not, as my simplistic introduction made it seem, that writing was a terrible new technology to be avoided at all costs. No, Plato wanted his students to understand that writing on its own was pointless: writing requires both a

thinker, to explain and defend and share their thoughts, and a reader, an audience to wrestle with, critique, and engage with those thoughts. It is a quintessentially social process and a quintessentially human one. To belabor the point, it requires humans, plural, and absolutely cannot be accomplished by a machine that cannot reason no matter how often and enthusiastically its creators insist it is intelligent. To believe otherwise, Plato insists, is "sowing them through a pen with words which cannot defend themselves by argument and cannot teach the truth effectually (Plato 1925, 276c). Our task now is to teach this same lesson to our students.

Publication Details and Disclosures

Acknowledgments

I would like to express my deepest appreciation to my colleagues in the Humanities Department at Fairmont State University for endless conversations about these issues, especially Dr. Donna Long and Dr. Kimberly Wickham, and my sincerest thanks to Dr. Angela Schwer for her constant support and mentorship. I first presented some of these ideas to my colleagues at Fairmont State at a conference on AI in the classroom and owe a debt of gratitude to Fairmont's Center for Teaching and Learning Innovation. I would also like to thank the attendees and organizers of the 2024 Teaching in the Age of AI conference at the University of Florida, Gainesville for their feedback and encouragement, as well as the opportunity to share the ideas that eventually coalesced into this article. Finally, my thanks to the kind and thoughtful reviewers of the early drafts at JWAI, who provided excellent advice for revision.

Funding

This project was not funded by any grants or specifically designated research funds. Conference travel funds from Fairmont State University's Humanities Department were used to help present the original ideas on which this project is based.

Generative AI Use

All of the writing in this text was composed by humans. I used Zotero, open-source software, to format the references.

Biography

I am an Assistant Professor of English at Fairmont State University. I teach introductory composition courses, the early British and World Literature surveys, and a variety of medieval literature courses. My research interests include first-year writing, medieval drama, and the ways that fiction both reflects and affects reality.

Copyright

References

ALTMAN, Sam (@sama), 2022. i am a stochastic parrot, and so r u. *X*. Online. 4 December 2022. Available from: https://x.com/sama/status/1599471830255177728

BENDER, Emily M. and HANNA, Alex, 2025. *The AI Con: How to Fight Big Tech's Hype and Create the Future We Want*. First edition. New York, NY: Harper.

BENDER, Emily M., GEBRU, Timnit, MCMILLAN-MAJOR, Angelina and SHMITCHELL, Shmargaret, 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜. In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. Online. Virtual Event Canada: ACM. 2021. p. 610–623. ISBN 9781450383097. DOI 10.1145/3442188.3445922.

COLWELL, Mary, 2015. How Jaws Misrepresented the Great White. *BBC News*. Online. 9 June 2015. Available from: https://www.bbc.com/news/magazine-33049099

DARBY, Flower, 2023. Why You Should Rethink Your Resistance to ChatGPT. *The Chronicle of Higher Education*. Online. 13 November 2023. Available from: https://www.chronicle.com/article/why-you-should-rethink-your-resistance-to-chatgpt

DUTTA, Arka, KHORRAMROUZ, Adel, DUTTA, Sujan and KHUDABUKHSH, Ashiqur R., 2024. Down the Toxicity Rabbit Hole: A Framework to Bias Audit Large Language Models with Key Emphasis on Racism, Antisemitism, and Misogyny. In: *Proceedings of the Thirty-ThirdInternational Joint Conference on Artificial Intelligence*. Online. Jeju, South Korea: International Joint Conferences on Artificial Intelligence Organization. 2024. p. 7242–7250. DOI 10.24963/ijcai.2024/801.

FOREIGN PRESS, 2023. A New Study's Surprising Insights into the Vulnerability of Gen Z and Millennials to Fake News. *Foreign Press*. Online. 16 July 2023. Available from: https://foreignpress.org/journalism-resources/a-new-studys-surprising-insights-into-the-vulnerability-of-gen-z-and-millennials-to-fake-news

GOODLAD, Lauren M. E. and BAKER, Samuel, 2023. Now the Humanities Can Disrupt "AI". *Public Books*. Online. 23 February 2023. Available from: https://www.publicbooks.org/now-the-humanities-can-disrupt-ai/

HARI, Johann, 2022. *Stolen Focus: Why You Can't Pay Attention—and How to Think Deeply Again*. New York: Crown.

HOGAN, Mél (@mel_hogan), 2024. There is no responsible or ethical use of AI when it comes to genAI. None. *X*. 20 September 2024.

KERR, Dara, 2024. How Memphis Became a Battleground over Elon Musk's xAI Supercomputer. *NPR*. Online. 11 September 2024. Available from: https://www.npr.org/2024/09/11/nx-s1-5088134/elon-musk-ai-xai-supercomputer-memphis-pollution

KILLEN, Kyle and KANE, Steven, 2022. *Halo*. 2022. New York: Showtime.

LATHAM, Scott, 2025. Are You Ready for the AI University?. *The Chronicle of Higher Education*. Online. 8 April 2025. Available from: https://www.chronicle.com/article/are-you-ready-for-the-ai-university

LI, Pengfei, YANG, Jianyi, WIERMAN, Adam and REN, Shaolei, 2024. Towards Environmentally Equitable AI via Geographical Load Balancing. In: *The 15th ACM International Conference on Future and Sustainable Energy Systems*. Online. Singapore Singapore: ACM. 31 May 2024. p. 291–307. DOI 10.1145/3632775.3661938.

MACLAREN, Iain (@ian_macl), 2024. Another course advertising 'ethical use of Gen AI in education'. How? Does ethics not include environmental impact, intellectual theft, bias, mis-selling, lying, lack of trust, etc? Or do we just talk about these things whilst still using it? #GenuineQuestion. *X*. 19 September 2024.

METZ, Cade, 2023. Chatbots May 'Hallucinate' More Often than Many Realize. *The New York Times*. Online. 16 November 2023. Available from: https://www.nytimes.com/2023/11/06/technology/chatbots-hallucination-rates.html

NEWCOMBE, Laura, 2025. Elon Musk's Grok AI Is as Gross as You Might Expect. *Gizmodo*. Online. 7 May 2025. Available from: https://gizmodo.com/elon-musks-grok-ai-is-as-gross-as-you-might-expect-2000599098

NICOLAS, Melissa, 2023. Eliminate the Required First-Year Writing Course. *Inside Higher Ed*. Online. 14 November 2023. Available from: https://www.insidehighered.com/opinion/views/2023/11/14/eliminate-required-first-year-writing-course-opinion

OPEN LETTER, 2015. Research Priorities for Robust and Beneficial Artificial Intelligence: An Open Letter. *Future of Life Institute*. Online. 28 October 2015. Available from: https://futureoflife.org/open-letter/ai-open-letter/

OPEN LETTER, 2023. Pause Giant AI Experiments: An Open Letter. *Future of Life Institute*. Online. 22 March 2023. Available from: https://futureoflife.org/open-letter/pause-giant-ai-experiments/

PADDISON, Laura, 2024. The System That Moves Water around the Earth Is Off Balance for the First Time in Human History. *CNN*. Online. 16 October 2024. Available from: https://www.cnn.com/2024/10/16/climate/global-water-cycle-off-balance-food-production

PLATO, 1925. *Phaedrus*. Online. Cambridge: Harvard University Press. Plato in Twelve Volumes. Available from: http://data.perseus.org/texts/urn:cts:greekLit:tlg0059.tlg012

SHELLEY, Mary, 1818. *Frankenstein; or, The Modern Prometheus*.

SPIELBERG, Stephen, 1975. *Jaws*. Universal City, CA: Universal.

TUFEKCI, Zeynep, 2025. For One Hilarious, Terrifying Day, Elon Musk's Chatbot Lost Its Mind. *The New York Times*. Online. 17 May 2025. Available from: https://www.nytimes.com/2025/05/17/opinion/grok-ai-musk-x-south-africa.html

WALSH, James D., 2025. Everyone Is Cheating Their Way Through College. *Intelligencer*. Online. 7 May 2025. Available from: https://nymag.com/intelligencer/article/openai-chatgpt-ai-cheating-education-college-students-school.html

WARNER, John, 2018. *Why They Can't Write: Killing the Five-Paragraph Essay and Other Necessities*. Baltimore, Maryland: Johns Hopkins University Press.

WARNER, John, 2019. *The Writer's Practice: Building Confidence in Your Nonfiction Writing*. New York: Penguin Publishing Group.

WARNER, John, 2025. *More Than Words: How to Think about Writing in the Age of AI*. First edition. New York: Basic Books.

WEIL, Elizabeth, 2023. You Are Not a Parrot. *Intelligencer*. Online. 2023. Available from: https://nymag.com/intelligencer/article/ai-artificial-intelligence-chatbots-emily-m-bender.html