

**Philosophical Warnings on the Use of AI in Education
Timely Advice from Plato, Martin Heidegger, Zhuangzi, and C. S. Lewis**

Dr. Mark J. BOONE

Hong Kong Baptist University, Hong Kong SAR, China
ORCID: [0000-0001-5766-6266](https://orcid.org/0000-0001-5766-6266)

Abstract: *Some important warnings about how we use technology in the philosophies of Plato, Martin Heidegger, Zhuangzi, and C. S. Lewis are relevant to the use of AI in education. Plato cautions us concerning what is lost when we let technology replace some of our own thinking processes. Far from making us more intelligent, the use of AI in writing falls into the mistakes Plato warns us against: We get lazy with learning and remembering, and we substitute a bundle of information for the wisdom and comprehension that constitute genuine knowledge. Heidegger advises against using technology to create more of a product and reducing the role of humanity to merely a part of the system of production. When writing with AI, we abandon our responsibility to shepherd our own work, and we become tools in the machinery of creating a written product, even letting the software guide us rather than the other way around. Zhuangzi teaches us not to follow the patterns set by social convention. Material written by AI is a distillation of conventional word-patterns. Lewis warns us that we abolish part of humanity when we use technology to get what we want without first learning to love what is good. Using AI to get our writing done means sacrificing the essential human love for finding and understanding the truth, instead allowing our own words to be conditioned by unknown authors of algorithms. In this article I explain these matters and close with some helpful suggestions for how we might use AI more constructively—assuming we are going to use it at all.*

Keywords: *Artificial intelligence, education, Plato, C. S. Lewis, Martin Heidegger, Zhuangzi, philosophy of education*

1. Introduction

The great philosophers have some interesting things to say about artificial intelligence (AI). Plato, Martin Heidegger, Zhuangzi, and C. S. Lewis are among them. They have some serious warnings about the use of AI in academic endeavors. *Plato* warns us that something is lost when we let some new form of technology do some of our thinking for us. This applies to the use of AI in writing, a practice which trains us to use preexisting patterns of words with minimal comprehension. *Heidegger* explains that when our overall aim is to produce more, we make ourselves cogs in a machine alongside other pieces of technology. The use of AI in writing has a similar effect, for, instead of being *people* who are guiding a piece of writing, we become a tool working with, and even being *guided by*, software as we work to produce a written product. *Zhuangzi* warns us against following the conventions of society when these are an imposition on nature,

freedom, and happiness. The use of AI in writing is *nothing but* a following of preexisting linguistic patterns society placed on the internet. When we let AI do our writing for us, we are following social convention — not patterns set by our own minds, and certainly not the best patterns, those of the Dao. *Lewis* explains that we go so far as to abolish humanity when we use technology to get what we want without any trained love of what's good. One important application to AI-writing is that we lose the essential human love of discovering and comprehending the truth when we substitute language patterns shaped by an algorithm-writer for our own understanding.

We will consider these matters one philosopher at a time.

2. Plato

Plato's *Phaedrus* explains that something is lost when we let some new form of technology do some of our work of thinking for us. Socrates tells a story about the god Theuth sharing his invention of writing with Thamus, the king of Egypt. (A recommended introduction to these themes in *Phaedrus* is Coblenz, 2016.) The *Phaedrus* must be quoted at length:

Theuth came to exhibit his arts to him and urged him to disseminate them to all the Egyptians. Thamus asked him about the usefulness of each art, and while Theuth was explaining it, Thamus praised him for whatever he thought was right in his explanations and criticized him for whatever he thought was wrong.

(...) But when they came to writing, Theuth said: "O King, here is something that, once learned, will make the Egyptians wiser and will improve their memory; I have discovered a potion for memory and for wisdom." Thamus, however, replied: "O most expert Theuth, one man can give birth to the elements of an art, but only another can judge how they can benefit or harm those who will use them. And now, since you are the father of writing, your affection for it has made you describe its effects as the opposite of what they really are. In fact, it will introduce forgetfulness into the soul of those who learn it: they will not practice using their memory because they will put their trust in writing, which is external and depends on signs that belong to others, instead of trying to remember from the inside, completely on their own. You have not discovered a potion for remembering, but for reminding; you provide your students with the appearance of wisdom, not with its reality. Your invention will enable them to hear many things without being properly taught, and they will imagine that they have come to know much while for the most part they will know nothing (...)"

Phaedrus, writing shares a strange feature with painting. The offsprings of painting stand there as if they are alive, but if anyone asks them anything, they remain most solemnly silent. The same is true of written words. You'd think they were speaking as if they had some understanding, but if you question anything that has been said because you want to learn more, it continues to signify just that very same thing forever. When once it has been written down, every discourse roams about everywhere, reaching indiscriminately those with understanding no less than those who have no business with it, and it doesn't know to whom it should speak and to whom it should not. (Plato, 1997, 551-552)

The lengthy passage from *Phaedrus* highlights key concerns raised by the king of Egypt regarding the technology of writing. He insightfully observes that writing encourages forgetfulness by diminishing the need for active memory recall. It does not *extend* memory; it *replaces* memory, and we get lazy and use writing instead. Another problem is that writing makes us think we have wisdom just because we have a lot of information. That information may have nothing to do with wisdom, and even if it does, we likely will not understand it. A third problem is that information travels everywhere, never coming to rest in the

minds of those who are capable of knowledge; this information is dead and without any comprehension. Socrates goes on to explain that true “discourse,” of which writing is a mere image, is written in the soul (Plato, 1997, 552-553). It “can defend itself, and it knows for whom it should speak and for whom it should remain silent.” It is, in Pheadrus’ summary, “the living, breathing discourse of the man who knows (...)” This discourse or word—the Greek *logos*—is the sort of thing a real philosopher, a seeker of wisdom, is after. It involves real understanding or knowledge—Greek *epistehmeh*—and so can respond to challenges.

Now Plato’s own writings are clearly meant to avoid the problems Socrates attributes to writing. They advocate that we seek wisdom, perhaps accepting a doctrine taught in the writing but certainly not *stopping* there—there is more to learn, and we don’t fully understand that doctrine yet anyway. Plato’s writings, through immersion in logical disputes, train us to not be satisfied with lifeless, mindless information but to seek the true *logos* of understanding of reality. (Of course, this would ultimately mean Plato’s teachings on non-physical reality.)

Now what does this have to do with AI? In general, *something is lost when we let some new form of technology do some of our thinking for us*. The specific problems observed by Socrates are *very* applicable to AI. AI lets us be lazy with what we learn and remember. AI tempts us to think we have some wisdom when all we have is a pile of information, and probably low-quality information at that. With AI, information spreads quickly, but it’s dead and mindless information neither leading to nor drawing from any conscious understanding.

Using writing instead of memorization was one technological step that gives something and takes something away, giving us the ability to organize massively larger amounts of information while keeping a good deal less of it in our heads. Modern typewriters brought writing that can be done more quickly and very readably, but our ability to use quality penmanship no doubt declined at the same time. Computers with word-processing software allowed us to write more quickly with better spelling while easily correcting our mistakes; but many of us no doubt have worse spelling due to our reliance on spelling-checking software, and we write more carelessly with disorganized thoughts because we think it will be easy to correct and reorganize later.

AI-writing is a massive step beyond this. Using AI in our writing can bring some benefits, such as the ability to produce (somewhat) reliable summaries of large blocks of text in an instant. But what is it, like all the other writing technologies, taking away? If we’re not careful, AI writing trains us to follow preexisting patterns of words with minimal comprehension.

Observations of student submissions reveal notable changes following the introduction of ChatGPT. Turnitin similarity scores have generally decreased, while the overall quality of English grammar in written assignments has improved. But I often find in student assignments piles of information consistent with easily accessible internet material, with little to no evidence of conscious interaction with course content. My students are accumulating information with no knowledge. (But more on that in good time.)

3. Heidegger

Heidegger is a 1900s German philosopher whose writings call us to be more attentive to the world and our presence in it—and, in particular, to how the two interact. The world is largely made by us, those strange creatures who live and work and think in it. Heidegger’s “Question Concerning Technology” is a classic work in the philosophy of science. Heidegger explains that, when we’re just trying to produce *more*, we make ourselves mere cogs in a machine, alongside the other pieces of technology that produce. AI-

writing makes us and the AI software just tools to produce the essay. We are not *people* putting ourselves into a piece of writing. The more positive philosophy of Heidegger is designed with humanity in mind—man or “Dasein,” the there-being, the being that is conscious of its own being and responsible for acting accordingly. We should do *that*, he thinks—we should live in a world conscious of and attentive to the world we are making. We should guide that world—shepherd it.

In *The Question Concerning Technology* (QCT), Heidegger introduces the concept of the 'standing-reserve' (German: Bestand), describing how technology transforms entities into resources available for immediate use. He asserts that 'everywhere everything is ordered to stand by, to be immediately on hand, indeed, to stand there just so that it may be on call for a further ordering' (Heidegger, 1977, p. 332). Technology involves an ordering of the world, transforming it—fields, rivers, trees, and cattle—into things standing by for our use when they are called for. Heidegger introduces the term *enframing* to describe the transformative process of the modern technological age, wherein humanity itself becomes integrated into technological systems (Heidegger, 1977, p. 325). At some risk of oversimplifying Heidegger: In a technological society, the world becomes nothing but standing reserve, and man pays a price for what he has made of the world; man “comes to the point where he himself will have to be taken as standing-reserve” (Heidegger, 1977, p. 332). The threat to humanity is not machines that kill; “The actual threat has already afflicted man in his essence” (Heidegger, 1977, p. 333). The threat to humanity is that we are reduced to nothing but standing-reserve—nothing but cogs in the machine into which we have shaped the world.

This concept applies to AI-generated writing by illustrating how humanity risks being reduced to standing-reserve when two conditions are met: prioritizing the production of academic writing as an end in itself and diminishing the human element to a mere component within the technological process of text creation. And this is what we do when we remove conscious human thought from a position of control over academic writing and surrender our decisions over what we say to AI.

Heidegger has a more positive vision for humanity, which we can find in his “Letter on Humanism” (Heidegger, 1977). Here, in an essay distinguishing himself from Sartre’s existentialist philosophy (in the latter’s famous “Existentialism Is a Humanism” talk), Heidegger considers what humanism is and what it requires. Our old metaphysical categories have failed, Heidegger thinks. Philosophers used to talk about Being, ignoring both the world in which humans live and the humans who live in it. The task for human beings is not to contemplate some higher reality, but to dwell in the reality we are in—and to *manage* it. We do not create this world, but we do have the ability to shed some light on it—and to, at least partially, determine what is the nature of this world we live in anyway. This is not a philosophy that is interested in whether a God exists beyond this world, but a philosophy interested in restoring a sense of divinity in everyday life. Heidegger comments on a line from the Greek philosopher Heraclitus: “Here too the gods come to presence” (Heidegger, 1977, p. 257). At the time, Heraclitus was warming himself by a fire and explaining himself to tourists who wanted to see the great thinker in action. Here, he explains, in the everyday world of man, divinity unfolds. Heidegger’s positive vision for this is subtle, and I cannot claim fully to understand it. But it does involve the responsibility of human beings to guide, cultivate, develop, or manage the world. “Man is not the lord of beings. Man is the shepherd of Being (...) Man is the neighbor of Being,” says Heidegger (Heidegger, 1977, p. 245). The world is a human phenomenon, though not one we entirely create or control. It is something we discover, to a certain degree, while also shaping, developing, and re-creating it in our own way (to a certain degree).

And how does this apply to AI writing? Quite simply, we cannot shepherd Being if we allow a machine to dictate how we think or what we say. We should not be mere tools in the creation of a written product.

We should be guiding the whole process, thinking independently of the machine—guiding *it*, not the other way around. Picture a scenario that I suspect is happening with some regularity in my own courses now: A student takes some information from the internet about Aristotle or Kant, perhaps supplementing it with some typed course notes, and runs them through ChatGPT, asking it to summarize the main ideas. The student then uses this information, perhaps with some minor re-writing, as her submission for a class assignment. The answer she submits may be essentially correct, but she has not pondered, sifted, and synthesized it for herself. She allowed the technology to shepherd her through the hardest and the most important work of thinking through it; all she did of her own agency was delegate this task to the program and then market the product to her teacher when it was finished. She has made herself a mere tool in the creation of a bit of writing she thinks her teacher will like—a tool sometimes controlled by another tool.

4. Zhuangzi

Zhuangzi, the ancient Daoist philosopher, is terribly subtle. Unless he's not. Maybe he's very simple and direct and we find him hard to understand because we are looking for something more complicated! In any case, I cannot pretend to have mastered Zhuangzi. But there are one or two things in Zhuangzi that I can understand well enough to relate to the use of AI in academic writing.

The idea I want to convey here is very simple: The use of AI in academic writing is nothing more than an exercise in following conventional patterns of thought and word and, as such, is not what we should be doing in academic writing. AI as we know it is, of course, *nothing but* an imitation of conventional patterns. It does not, strictly speaking, imitate conventional *thoughts*, but that is only because it does not *have* any thoughts; if it did, it would be imitating conventional ones. If we read what it says and agree with it, that is what *we* are doing. And if we put our name on AI-writing, then we are putting our name on conventional patterns of thought.

Perhaps, before getting to Zhuangzi himself, we should look at some examples. Here is a sample of letting conventional patterns give us the answers:

Wu-wei is an idea that originated in Daoism, an ancient branch of Chinese philosophy. Translations of wu-wei include “non-action” and “effortless action.” It is frequently defined as a condition of being in which one naturally and without effort or resistance links with the universe's natural flow.

It does not mean total passivity or inactivity in the context of wu-wei. In fact, it offers an approach to living in balance with the natural order of things, free from needless effort or conflict. It entails behaving impulsively and intuitively while letting go of one's self-image or choice.

A student submitted this in an online weekly reflection assignment in the spring semester of 2024. I am pretty sure it was written by ChatGPT. The answer may be more or less correct, but the student probably learned nothing, and what really bothers me is there is not even a hint of conscious interaction with course content, which is why I gave this a bad grade. Ironically, the student's work here somewhat resembles Daoist *wu-wei* in that it was not very hard work, but it completely fails to be like what it describes since the student himself is not even involved in the answer.

One student submitted this concerning the *Analects*:

“I agree with the emphasis on cultivating moral character, practicing benevolence, and pursuing knowledge aligns with the idea of personal growth and ethical conduct. Additionally, the promotion of

harmonious relationships, respect for elders, and the pursuit of balance and moderation can contribute to a stable and cohesive society. These teachings reflect values that many individuals and cultures recognize as important for personal and social well-being”.

The first sentence is a run-on, so maybe the student wrote it! I have difficulty imagining any undergrad writing the second and third sentences. My grading standard was not that I was sure this was ChatGPT, but that (again) there was no hint of any conscious interaction with course content. Students who interact with course content normally express their thoughts on concepts such as the virtue of Ren (benevolence/humanity) or the Golden Rule—topics I address in my lectures—instead of using undefined terms like “pursuit of balance.” Most interesting, however, is the total abandonment of any responsibility to think for oneself: Why should I or the student care about what “many individuals and cultures” think is important? The student should have explained what *the student* thinks is important, and *why*.

I will let one longer example pass with little comment. The question was “What does Nietzsche say about Christianity and otherworldliness?” The student’s answer recycles, with minimal clarity, a lot of general information available online and gives no hint of interacting either with the reading or with the simple explanation in class notes. The unusually good grammar and American spelling (in Hong Kong!) only help to confirm that the student did little work and learned little:

Nietzsche argued that Christianity had a negative impact on human existence because it emphasized otherworldly ideals and the afterlife. According to him, Christianity promotes rejecting life and depreciating the present in favor of an idealized eternity. According to Nietzsche, this concentration on the hereafter devalued the present moment and rejected human inclinations and aspirations.

Nietzsche stated that Christianity’s reliance on a hereafter fostered a morality of slavery. He claimed that Christianity’s moral framework, which emphasized meekness, humility, and self-sacrifice, suppressed the intrinsic drives for dominance, power, and self-assertion. This, Nietzsche believes, stops people from leading satisfying lives and becoming powerful, morally upright individuals.

Nietzsche also criticized Christianity for fostering herd behavior and diminishing individualism. He believed that Christianity hampered personal development and the pursuit of individual greatness because it fostered conformity and discouraged critical thinking.

Nietzsche’s critique of Christianity was part of a broader critique of traditional morality and ethics. He sought to disrupt the mainstream religious and moral systems by encouraging a re-evaluation of values and the development of a new moral framework that embraced life-affirming attributes such as strength, self-assertion, and creativity.

It is critical to recall that Nietzsche held nuanced and complex views on Christianity and otherworldliness, which should be understood considering his larger philosophical concepts. His criticisms of Christianity included cultural and societal implications, as well as theological and ecclesiastical concerns.

Why would it be considered appropriate to let a piece of technology reconstruct *everyone else’s* thought patterns in a person’s *own* writings?

Let us turn our attention to Zhuangzi himself. What remains unclear to me is the specific patterns that should guide our thoughts. It seems reasonable to suggest “Follow the facts!” or “Follow the way things are!” or “Think for yourself!” or “Follow the Dao!” (The *Dao* is the Way—the pattern of all that is, and the

pattern by which we should live.) All these simple statements, probably, capture one aspect of Zhuangzi's thought while failing adequately to explain it.

What I *can* state with some degree of clarity is the negative: Zhuangzi clearly says that we should not, for Heaven's sake, follow the patterns of conventional thought! As Watson summarizes, the lesson of *Zhuangzi* is "free yourself from the world" (Watson, 2003, p. 3). Again, the goal of this philosophy is to become one "who has freed himself from conventional standards of judgment . . ." (Watson, 2003, p. 5). In some sayings in *Zhuangzi* which are attributed to Confucius, we read, "Just go along with things and let your mind move freely. Resign yourself to what cannot be avoided and nourish what is within you—that is best. What more do you have to do to fulfill your mission? Nothing is as good as following orders (obeying fate)—that's how difficult it is!" (Zhuangzi, 2003, p. 57). Watson explains in a footnote that the alternative translations "following orders" and "obeying fate" are *both* "almost certainly intended," and that the statement of difficulty is ironic. It is *easy* to obey fate. The clearer lesson is to let one's mind freely move while facing facts and cultivating oneself. Following conventional social rules is not recommended and is implicitly excluded.

Another attribution to Confucius lists "Life, death, preservation, loss, failure, success, poverty, riches, worthiness, unworthiness, slander, fame, hunger, thirst, cold, heat" as "alternations of the world, the workings of fate" (Zhuangzi, 2003, p.69). They come and go in cycles. As *Daodejing* says, "Misfortune depends upon fortune. Fortune conceals misfortune" (Laozi, 1991, section 58). Zhuangzi's Confucius continues: One should "harmonize and delight in them, master them and never be at a loss for joy." These are the world's conventional priorities, but they are the wrong ones. Wisdom thinks independently of them. (Similarly, Kipling's poem "If" urges us to recognize both Triumph and Disaster as imposters, the teacher of Ecclesiastes tells us that there are a time and a place for everything under the sun, Boethius reminds us of Fortune's constantly turning wheel, and so on in other great works of wisdom literature.)

In one subtle passage, Zhuangzi has the god of the North Sea talking to the god of the Yellow River. The former explains that "He who understands the Way is certain to have command of basic principles," and thus can deal with whatever life throws at him (Zhuangzi, 2003, p.105). He is stable and contented in any circumstances. (Again, we also think of western sources, such as Boethius and the Stoics.) His state of being is based on what is heavenly (or, perhaps, on the natural), like the fact that horses have four feet. It is not based on human intervention, like putting a halter on a horse, and does not allow human intervention to undermine what is heavenly (or natural). No social convention would replace such a person's understanding of things.

Zhuangzi advises: "Embody to the fullest what has no end and wander where there is no trail (...). The Perfect Man uses his mind like a mirror—going after nothing, welcoming nothing, responding but not storing" (Zhuangzi, 2003, p. 95). The Daoist sage is independent of social priorities and thinks independently of social rules for how to think.

But, of course, AI writing is nothing more than replicating the internet's conventions for how and what to think. As Zhuangzi says, "Words exist because of meaning; once you've gotten the meaning, you can forget the words. Where can I find a man who has forgotten words so I can have a word with him?" (Zhuangzi, 2003, p.141). There is a time to write nothing at all, understanding a meaning without words. But there is never a time to write words while trusting on social conventions to supply a meaning distilled by a program with no awareness of that meaning.

let's defer to a scholar of Zhuangzi for the final insight:

We ordinarily confine ourselves within our social roles, expectations, and values, and with our everyday understandings of things. But this, according to Zhuangzi, is inadequate for a deeper appreciation of the natures of things, and for a more successful mode of interacting with them. We need at the very least to undo preconceptions that prevent us from seeing things and events in new ways; we need to see how we can structure and restructure the boundaries of things. (Coutinho, N.d., Section 3.a)

An important follow-up question arises: How could we possibly hope to get past social conventions and undo our preconceptions if our writing consists of a distillation of them? My students' work, copied above, exemplifies this mistake. It distills conventional observations on some philosophers instead of thinking through the writings for oneself.

5. Lewis

We will close with *The Abolition of Man*, a modern classic in the philosophy of technology and an important book in philosophy of education and in ethics. We will briefly consider some of the main points relevant to AI writing. (For a more adequate introduction to *Abolition*, see Boone, 2016.) Lewis warns that a human being, as traditionally understood by all prior philosophies and civilizations, recognizes that some things are objectively valuable, and is supposed to respond to this value. But modern culture has embarked on a project of debunking values, the end-result of which is terrifying: If some people who reject the teaching that there are objective values should but once possess the technological power to reshape humanity according to their whims, they will surely do so. AI-writing is not, in and of itself, the fulfillment of Lewis' disturbing vision; but it does partake of it just a little. In place of cultivating the love of truth, when we engage in AI-writing we are submitting our own understanding of the truth to word-patterns produced by algorithms and guided by the preferences of the people who wrote them.

Abolition begins by observing that English textbooks are teaching philosophy, but surreptitiously (Lewis, 1947, chapter 1). Under the pretense of teaching students how to evaluate texts, they proclaim that there is no such thing as a factual evaluation statement, but only a personal (and subjective) emotional preference. Lewis explains that this is a radical philosophy. (It's also a *bad* philosophy. It's the philosophy of the worst chapter in the entire history of western philosophy, Logical Positivism. But that is a story for another time; see Boone, 2016, pp. 6-8.) This philosophy completely rejects the traditional view shared by all prior philosophies and civilizations: that some things are objectively right or wrong. Lewis further explains that on this traditional philosophy what it means to be *human* is to have a correct response to objective value. This is the teaching of the Tao, as he calls it, borrowing Confucian and Daoist language. (Or the *Dao*. Lewis uses "Tao," based on an older transliteration system than the now-dominant system which gives us "Dao.") This response to objective value is mental, but at least as much emotional as mental: We ought to enjoy the company of cute little children, be humbled by the venerability of old men, and be awed by the sublimity of a breathtaking waterfall. Lewis further explains that there is no alternative to recognizing that some values are fundamental (Lewis, 1947, chapter 2). We cannot consistently debunk one value on the grounds that values are subjective and yet affirm another value—not even the value of knowing the truth and exposing spurious subjective claims! The only way to value anything at all is to recognize that *some* value simply *is real* independently of human perception. The alternative to this teaching is the Nietzschean one (Lewis, 1947, chapter 3). And that is to reject all value claims and simply do what we want because we want it.

Lewis warns that this approach will involve an abolition of humanity. That is, it will lead to the abolition of humanity *if* scientific know-how and technological power are up to the task. Prior civilizations understood a human being as a being having the capability of responding to objective facts about what is

valuable, and the responsibility to do it rightly. When we use technology to get what we want without any trained love of what is good, we are replacing such beings with something else entirely: These valueless techno-masters are recreating humanity simply according to their own preferences, for, without a recognition of values, some subjective preference is all they have left.

What does this have to do with AI writing? Quite simply: AI-writing displays just a bit of the folly of which Lewis warns. One essential trait of humanity—and one of the most important goals of education—is our love of discovering and comprehending the truth. When we use AI software to substitute language patterns for our own understanding, we are failing to cultivate that trait in ourselves. And we are replacing it with the trait of submission to the *algorithms* used by AI, and to the preferences of whoever wrote those algorithms. We have given up one aspect of our humanity, replacing it with conditioning by someone else.

6. Conclusion

Plato teaches us not to let technology substitute for the necessary work of human thought, producing more information with less comprehension. Students daily are using AI-writing to produce more information faster, but with less input from their own minds and less understanding. *Heidegger* warns us to not reduce ourselves to parts of a machine creating products. Academic writing should be a process of people using their writing tools to shape a little bit of the world. It should not be a process that uses people to create a product. Students are now letting AI create their writing, surrendering to an algorithm their own role as shepherds of the process. *Zhuangzi* urges us to not follow social convention, but AI-writing all too often takes social conventions as manifested on the internet and distills them into a neat little package. And *Lewis* says that we destroy our own humanity when we use powerful technology to get what we want without first learning to want what is good. AI-writing sacrifices the human love of discovering and understanding the truth, submitting our words to conditioning by someone who wrote some algorithms.

These great philosophies warn us that the use of AI in writing can steal a piece of our humanity from us. A character in *Star Trek: Insurrection*, probably the most Daoist of Star Trek productions, introduced the idea well enough: “We believe when you create a machine to do the work of a man, you take something away from the man” (Frakes, 1999). (Perhaps this is overstated, but more on that in my final paragraph.)

How should we proceed with this information? The first step is to develop an awareness of the situation. Perhaps the strongest protection against losing some of ourselves to a careless use of this new technology is simply to recognize the value—the *objective* value, *Lewis* reminds us—of what is at risk. A human being, and some hard work by that human being in learning for himself to think and to communicate well—those things *matter*.

What steps should we take beyond this? Should the use of AI in academic writing be prohibited? That perspective is understandable: We didn’t need it in my day, I still have never used it in my writing, and I think any advantages it may present are not worth the risk.

But is this feasible? Likely not. AI in academic writing is probably here to stay for the foreseeable future. Moreover, it was never my intention to argue that no technology can enhance learning; I daresay sometimes it can. I have philosophical warnings concerning AI in education, not a moral law forbidding it.

What steps should we take moving forward? Well, now we cautiously move along with our lives with these things in mind. But if you want a concrete step, we can all take, I can suggest two things that make sense,

one for teaching and one for scholarly research. To present them in Heideggerian language, *man is the shepherd of being*, and we human beings need to be the shepherds of our academic writing.

When it comes to teaching our students, that means we mustn't let them rely on AI scrolling the internet to pile up a glittering heap of words. In the humanities at least, assignments should be focused on original sources, and students should not even be capable of finding decent answers on the internet. Questions should rely as much as possible on class readings, perhaps supplemented by lecture notes.

When it comes to doing our academic research, maybe there are some ways we can use AI like a staff in our hands to herd the sheep but without letting it lead us around. We should use AI like we use (or like we *should* be using) spell-checking software and search engines: It helps us out, prevents a few errors, and frees up a few brain cells for doing other things, but we sure as heck don't want to forget how to spell or how to look up books on library shelves.

If such metaphors and comparisons are not clear enough, some examples should help. Letting AI discover my thesis statement for me—that's *bad*. Having AI rephrase a tricky paragraph of my own composition to help me get past a moment of writer's block, and then carefully reading the result to make sure I agree with what it says—that's probably ok. Telling AI to write up a relevant bibliography and looking over the results to find out which sources are real, good, and relevant to my research so I can then go and cite them myself—that's probably ok. Having AI read an article and summarize it for me, so I don't have to think about it—that's *bad*. But having AI read and summarize it as an intro to the article to be handled as I might the work of a competing scholar with whom I often disagree—a useful intro in which I am keeping a sharp eye out for errors—that's probably ok.

A new technology may not be necessarily and inherently evil. But the use we make of it could be devastatingly wrong. *I'm not saying you can't let it help you do your work; but don't let a machine do the work of a human being*. May there be more reflection and dialogue about AI and education, and may we keep these things in mind as we continue this important discussion.

References

- Boethius. (n.d.). *The Consolation of Philosophy*. Penguin Classics.
- Boone, M. J. (2016). Introduction: Finding C. S. Lewis in science fiction film and television. In K. C. Neece & M. J. Boone (Eds.), *Science fiction film and the abolition of man: Finding C. S. Lewis in science fiction films and television* (pp. 1–21). Pickwick Publications.
- Coblentz, T. (2016). A vision of transcendence: Monstrous intelligence and loving understanding. In M. J. Boone & K. C. Neece (Eds.), *Science fiction and the abolition of man: Finding C. S. Lewis in science fiction films and television* (pp. 104–120). Pickwick Publications.
- Coutinho, S. (n.d.). Zhuangzi (Chuang-Tzu, 369–298 B.C.E.). *Internet Encyclopedia of Philosophy*. <https://iep.utm.edu/zhuangzi-chuang-tzu-chinese-philosopher/>
- Frakes, J. (Director). (1999). *Star Trek: Insurrection*. Paramount.
- Heidegger, M. (1977). Letter on humanism. In D. F. Krell (Ed.), *Martin Heidegger: Basic writings* (pp. 217–265). Harper & Row.
- Heidegger, M. (1977). The question concerning technology. In D. F. Krell (Ed.), *Martin Heidegger: Basic writings* (pp. 311–341). Harper & Row.
- Kipling, R. (n.d.). If. Retrieved from: <https://poets.org/poem/if>

-
- Laozi. (1991). *Daodejing* (A. C. Muller, Trans.). <http://www.acmuller.net/con-dao/daodejing.html>
- Lewis, C. S. (1947). *The abolition of man*. Macmillan.
- Plato. (1997). *Phaedrus* (A. Nehamas & P. Woodruff, Trans.). In J. M. Cooper (Ed.), *Plato: Complete works*. Hackett.
- Watson, B. (2003). Introduction to Zhuangzi: Basic writings. In B. Watson (Trans.), *Zhuangzi: Basic writings* (pp. vii–xv). Columbia University Press.
- Zhuangzi. (2003). *Zhuangzi: Basic writings* (B. Watson, Trans.). Columbia University Press.

Bionote:

Mark J. Boone is a philosopher specializing in the study of the great texts. Having taught in the USA and Pakistan, he is currently teaching at Hong Kong Baptist University. He can be contacted via: platoandaugustine@gmail.com.