Post-Process but Not Post-Writing: Large Language Models and a Future for Composition Pedagogy

S. Scott Graham

[S]tarting with a blurry copy of unoriginal work isn't a good way to create original work. If you're a writer, you will write a lot of unoriginal work before you write something original. And the time and effort expended on that unoriginal work isn't wasted; on the contrary, I would suggest that it is precisely what enables you to eventually create something original. The hours spent choosing the right word and rearranging sentences to better follow one another are what teach you how meaning is conveyed by prose… If students never have to write essays that we have all read before, they will never gain the skills needed to write something that we have never read.

—Ted Chiang

The above passage from Ted Chiang's *New Yorker* essay on ChatGPT has become stock-in-trade for critiques of using AI in student writing and writing pedagogy. Indeed, the internet now abounds with confident complaints about the damage ChatGPT and other AI systems will invariably do to student writing. These new technologies have subsequently been heralded as the end of the essay and a substantial threat to writing education (Marche). The common arguments about potential student use of these technologies suggest AI-assisted writing will damage student learning by shortcutting the writing process. They further worry that AI-based pedagogy will de-skill students by reducing writing to a mere editing practice. Additionally, some suggest that students will fail to do adequate background research, relying on dubious or non-existent AI-generated sources. Ultimately, these concerns indicate a view of AI-assisted writing where students can essentially parachute their way into the last stages of the writing process, skipping past considerable effort and learning. I largely reject these worries because they simply don't fit with my understanding of how writing works. At their core, each of these concerns is marked by a problematic commitment to a limited and linear process-based model of writing. Instead, I think writing instructors and scholars of writing studies can consider how our hard-won insights about diverse and variable writing practices might allow for robust writing pedagogies that engage, rather than prohibit, use of AI.

While engaging with AI may be uncomfortable for many, I think it should be considered a requirement for two reasons. First, the history of writing stud-

ies is a history of moral panic about new technologies. From Plato's laments about inscription itself to later rejections of erasers, typewriters, word processors, SMS, and spell-check as inimical to writing, writing studies was always already worried about the effects of new technologies. However, in each of these cases, we learned how to adapt our pedagogy to new realities and the end results have been that students now write more than ever before without any measurable decline in quality (Lunsford and Lunsford). Secondly, prohibition has always been doomed to failure. While I genuinely wish we lived in a society that thought carefully and strategically about adopting new technologies, we do not. To teach students about writing without engaging the writing technologies they have access to and will invariably use does them a substantial disservice. It reminds me of the early days of the word processor when professors who had learned to write in an earlier era would compel submission of hand-written drafts because they knew (they just knew) that pencils produce better writing and thinking. If you agree with me that we don't want to be the AI-equivalent of the pencil mandator, then we need a robust foundation for our AI writing efforts. The good news is that we already have one in post-process pedagogy; that's the main argument I will be making in this short essay. However, before I fully elucidate this argument, I'd like to step back and talk about what exactly these new AI systems can do.

What Even Are These Things?

Large Language Models (LLMs) like ChatGPT, GPT-{pick a number}, Galactica, LLaMA, and Bard generally work by vacuuming up giant piles of text available on the internet. ElutherAI is one of the few LLM companies out there that offer good documentation about the data they use. A review of their training data, "The Pile," shows a mix of PubMed Central articles, open web crawls, Reddit forums, Wikipedia, The Free Law project and so on (Gaoet al). With these types of data on board, LLMs transform the language of online text into mathematical representations of meaning and do a bunch of fancy math to learn what meanings usually follow from prior meanings (Bender et al). Put simply, an understanding of common linguistic probability distributions lets the model know that when finishing off the sentence, "The cat is wearing a…," "hat" is far more likely to be the next word than "ham sandwich."

With these large datasets and underlying probability distributions, LLMs produce text in response to users' prompts. By way of illustration, I put that last italicized sentence into ChatGPT (with no further information or instructions), and it followed up by telling me:

Yes, that's correct. Language models like GPT-3 are trained on large datasets of text, which allows them to learn the underlying probability distributions of language. This means that when you give the model a prompt or a starting sentence, it can use that knowledge to generate a continuation or a full text that is statistically likely based on what it has learned from the dataset. The model uses complex algorithms and calculations to determine the likelihood of each possible continuation, based on the probabilities it has learned from the data. This process is often referred to as autoregression, because the model generates each new word or phrase based on the words that came before it in the text. While LLMs like GPT-3 can produce impressive and even convincing responses to user prompts, it's important to remember that they are ultimately just algorithms that are generating text based on statistical patterns. They don't have true understanding or creativity like humans do, and their outputs can sometimes be nonsensical or even offensive if the training data contained biased or problematic language.

If you're just looking at examples like this, it might seem that LLMs are an easy pathway to decent writing, but that's just not the case. I like to think of LLMs as genre-singling engines. They use probability distributions to loudly indicate the nature of the output. But that output must reflect the input, and if your input doesn't also loudly signal the desired genre, the output will not be what you were looking for. What's now known as "prompt engineering" becomes a new and critical writing-related skill, because iterative prompt generation and output creation is essential to getting good content. What's more, even well-curated prompts may lead writers astray. The example above is all very good information, but as ChatGPT, itself, reminds us, it's dangerous to trust what these models say as they frequently produce incorrect information that happens to conform to the observed linguistic probabilities. Hallucinations has become the term-of-art for LLMs' propensity to make shit up, and it's a pervasive problem that requires diligent fact-checking to overcome. Ultimately, these impressive, but limited, tools have the potential to transform writing and writing practices for both good and ill. The ills side of things gets a lot of airtime these days, so what I want to spend remainder of this short essay reflecting on one possible framework that supports robust use of AI as a part of writing pedagogy.

Theoretical Throwback

Unfortunately, I can no longer be confused for being among the youngest members of any given faculty. I received my PhD in 2010 from the program

in Rhetoric and Professional Communication at Iowa State University. In that place and time, writing pedagogy was essentially post-process pedagogy. The insights of Lee-Ann Kastman Breuch, Thomas Kent, David Wallace, Irene Ward, and others served as a center of gravity for understanding writing and writing practices. Ultimately, the post-process pedagogy of the era was a wholesale rejection of any canonical writing process. As far back as 1994, Lisa Ede noted that the writing process has been thoroughly "co-opted and commodified—by textbooks that oversimplified and rigidified a complex phenomenon, by overzealous language arts coordinators and writing program administrators who assumed that the process approach to teaching could be 'taught' in one or two in-service sessions" (35-36).

Sadly, little has changed in the early days of LLMs. If you Google "the writing process," you'll find a million-and-one variants of "the" writing process. Countless webpages, textbooks, and writing guides of all media replicate and propagate a sequential writing process that moves, generally, from ideation and pre-writing (notetaking, outlining) through drafting and revision to submission and publication. Despite the widespread popularity of this linear writing process model, as I note above, it has, of course, been long rejected by writing studies. Ultimately, the now old-school post-process approach may be best represented by Kastman Breuch's philosophical rumination on the nature of post-process pedagogy in the sadly moribund *JAC*. There, Kastman Breuch defined an approach to writing centered in recursive and dialogic processes of meaning-making. Building on insights from Kent and Ward, she advanced a notion of writing as reciprocal hermeneutic process that proceeds through iterative dialogic engagement with imagined audiences, teachers, mentors, gatekeepers, institutions, (peer) reviewers, and ideas (126). Figure 1 (below) is my distillation of the primary insights of post-process pedagogy.)
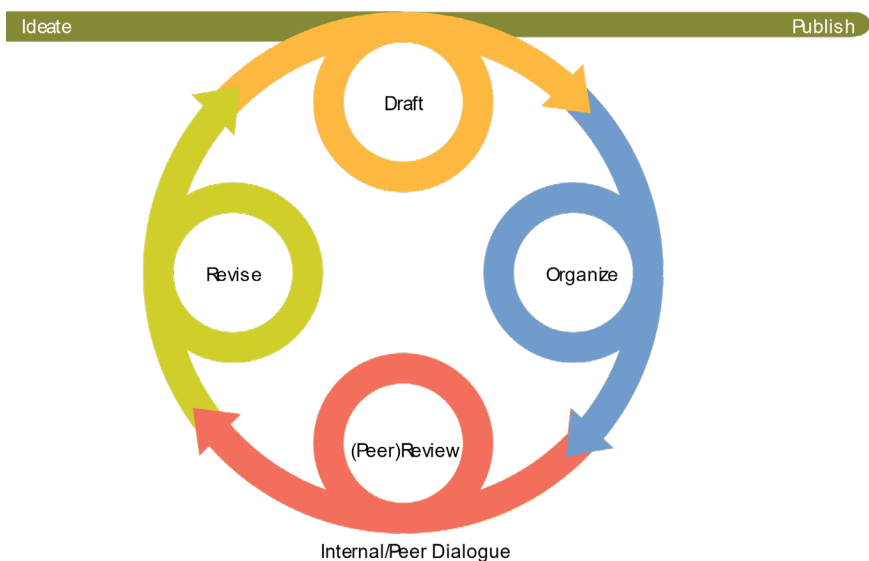
Draft

Revise

Organize

(Peer)Review

Internal/Peer Dialogue

Figure 1: A writing process defined by iterative recursive and dialogic steps.

Multidimensional Recursion

If we reject the simple linearity of the typical model, we have at least one alternative model for imagining where and how AI might slot into a robust writing pedagogy, and AI use no longer needs to be considered a matter of parachuting in at the last moment a pre-defined sequential writing process. Rather, AI provides the opportunity to add multiple dimensions of recursion where prompt-engineering, output curation, fact-checking, and revision become an orthogonal dimension to traditional writing and learning processes. The internal, peer, instructor, institutional, and social dialogues that support traditional writing practices are augmented through dialogic engagement with the LLM and its outputs (see figure 2). In the end, we have something that functions strikingly like Chiang's vision for good writing practices, one where re-writing and revision is essential. That is, writers still get "hours spent choosing the right word and rearranging sentences to better follow one another are what teach you how meaning is conveyed by prose" (Chiang). And, of course, this fits with already understood best practices in writing studies. When writing education is grounded in the best insights of writing studies, it aims principally at helping students develop more robust and recursive revision practices. This is part of the reason Ann Lamott's "Shitty First Drafts" is such a popular assigned reading in writing education: It's a colorful way to encourage students to get literally any text on paper so that they can actually start practicing the more robust and recursive revision/drafting practices

that lead to high quality writing. Well, you know what produces the kind of immediate mediocre quality and unoriginal text that's ideal for revision pedagogy? LLMs!
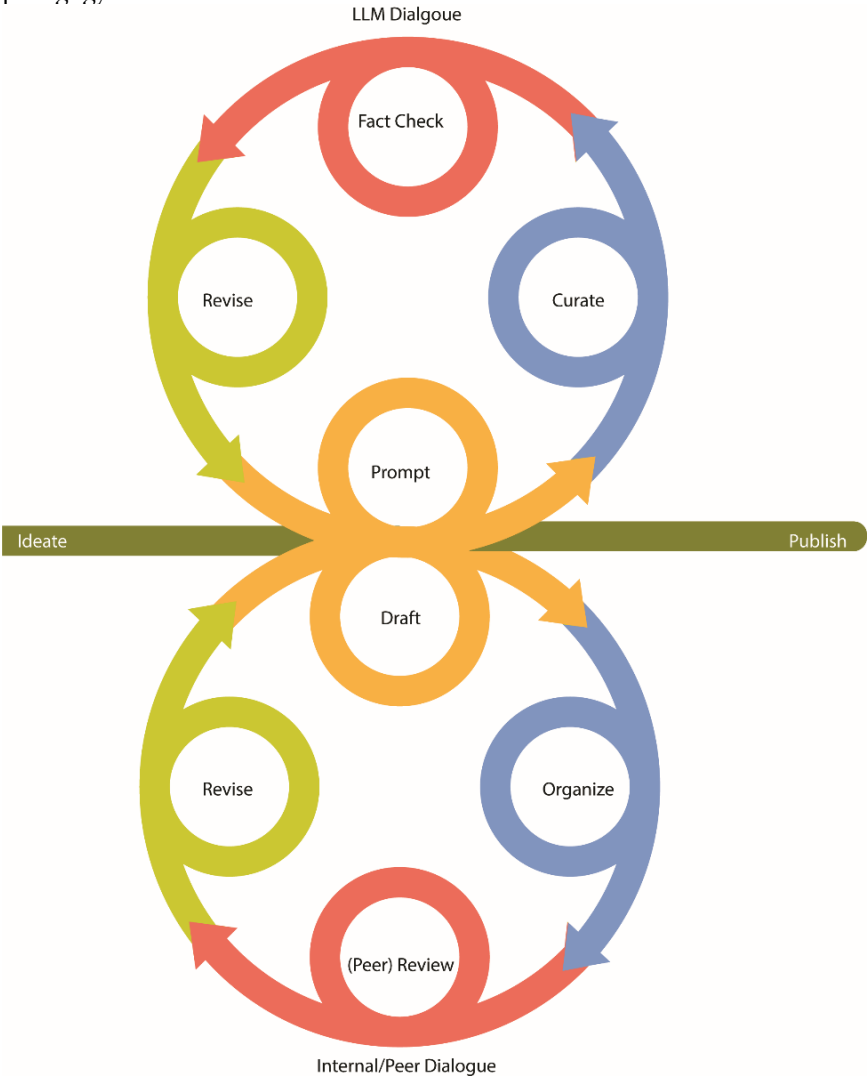

Figure 2: A multidimensional recursive AI-assisted writing process.

To be entirely fair, I'm not saying AI-assisted writing instruction grounded in post-process pedagogy will be easy to accomplish. It won't happen by accident, but good writing education never does. What we need going forward is a solid practical foundation and some rigorous empirical investigation. Writing studies should be taking the lead on developing prompt engineering modules,

new approaches to research and fact-checking pedagogy, rapid genre generation exercises, and output revision curricula. With innovative new practices in place, we can test what works and what doesn't. We can carefully evaluate which new pedagogies preserve learning and prepare students effectively for their future lives. Accomplishing this in the classroom will require carefully-crafted student learning activities to be sure. If we, as a discipline, opt out of this work, we'll both shortchange our students and cede critical ground to educational researchers in other disciplines who do not share our understanding of good writing and good writing pedagogy.

Works Cited

"With these large datasets and underlying probability distributions, LLMs produce text in response to users' prompts." prompt. *ChatGPT*, 23 Mar. version, OpenAI, 4 Apr. 2023, chat.openai.com/chat.

Bender, Emily M., et al. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. 2021, pp. 610–23, doi.org/10.1145/3442188.3445922.

Chiang, Ted. "ChatGPT Is a Blurry JPEG of the Web." *The New Yorker*, 9 Feb. 2023. www.newyorker.com/tech/annals-of-technology/chatgpt-is-a-blurry-jpeg-of-the-web. Accessed 4 Apr., 2023.

Ede, Lisa. "Reading the Writing Process." *Taking Stock: The Writing Process Movement in the '90s.* Eds. Lad Tobin and Thomas Newkirk. Boynton, 1994, pp. 31–43.

Gao, Leo et al. "The pile: An 800gb Dataset of Diverse Text for Language Modeling." *arXiv* preprint arXiv:2101.00027, 2020, doi.org/10.48550/arXiv.2101.00027.

Lamott, Anne. "Shitty First Drafts." *Writing about Writing: A College Reader.* ed. Elizabeth Wardle and Doug Downs, 1st ed. Bedford/St. Martins*, 2010, pp. 527–31.

Lunsford, Andrea A., and Karen J. Lunsford. "'Mistakes Are a Fact of Life': A National Comparative Study." *College Composition and Communication*, vol. 59, no. 4, 2008, pp. 781-806.

Marche, Stephen. "The College Essay Is Dead." *The Atlantic*, 6 Dec. 2022, www.theatlantic.com/technology/archive/2022/12/chatgpt-ai-writing-college-student-essays/672371/. Accessed 4 Apr., 2023.