

CHAPTER 20.

AI WRITING TOOLS CAN
“THINK” ✦ *HUMAN WRITERS ARE
ALWAYS THE REAL “THINKERS”*

Alex Helberg

Trinity College

If you’ve taken a class that teaches writing in high school or college, you’ve likely been told at some point that “writing is a form of thinking.” This truism has long been at the center of secondary and undergraduate writing pedagogy; it illustrates that the practice of writing provides us with an instrumentally useful method for representing complex ideas, testing them out in language, and shaping them to better engage with the thoughts of our readers.

If so many agree that “writing is a form of thinking,” it should not come as a great surprise that the generative artificial intelligence (GenAI) hype cycle that began in late 2022 hastened some to herald the arrival of “thinking machines”—if these tools can “write,” then surely they can also think! In American news media and education discourse, this has led to a much more problematic idea: that GenAI chatbots’ “writing” ability is evidence that the programs are not only modeling, but optimizing heretofore unseen processes of human cognition. One need look no further than headlines proclaiming that GenAI chatbots have proved their thinking prowess by “pass[ing] exams from law and business schools” (Kelly, 2023) or prognostications that advanced models “will have ‘Ph.D.-level’ intelligence” (Larsen, 2024) to see the assumption at play: language-generation is the surest sign of advanced cognitive capabilities for GenAI tools.

This assumption reflects what Alexis Baria and Keith Cross (2021) call the “Computational Metaphor” of cognition: the idea that human brains are mere information-processing hardware, and that computer hardware, by proxy, acts like a human brain. If we accede to this notion, we might be led to conclude that advanced machines can do our best thinking and communicating for us, rendering students and teachers of writing as little more than imperfect conduits for “clear” and “effective” language.

This idea that GenAI writing tools can *think* because they can “write” obscures the observable fact that these programs are not, in plain truth, actually writing or thinking the way that human writers do. Over the past several decades, writing

researchers have been studying the thinking processes that take place when we write, and have perennially returned to two key insights that all but preclude the possibility that current GenAI technologies can autonomously “think” because they can generate text:

- Writing is a non-linear process that requires recursive self-regulation.
- Sophisticated writing requires metacognitive reflection on the writers’ own process.

More plainly, it is still *human* thinking that propels AI text-generation. Rather than attributing cognition to machines, we would do better to examine the ways that human thinkers and writers *engage* with language models to produce text that mimics the product of human thinking.

WRITING IS NON-LINEAR AND REQUIRES RECURSIVE SELF-REGULATION

In the 1980s, the foundational work of Linda Flower and John R. Hayes (1981) focused researchers’ attention on the various cognitive processes that writers undergo when approaching a writing task. In their analysis of writers’ think-aloud protocols conducted throughout the different stages of a writing process, Flower and Hayes found that writers’ internal representation of the task—their purpose and goals, their audience, and other contextual factors influencing what they want to communicate—is central in guiding them through their composition process. Critically, they also found that these internal representations of goals and other knowledge that writers bring to their process are subject to “recursivity”—writers consistently revise and re-conceptualize their goals and the strategies they use to achieve them as the experience of writing unfolds in the moment. Recursivity, they emphasize, implies a lack of “linear” structure in how writers accomplish tasks, in that “a given process may be called upon at any time and embedded within another process or even within another instance of itself” (1981, p. 375). For example, a writer might begin an essay with her introduction as a way of getting the “big picture” representation of her argument onto the page, but then return and completely reconstruct this paragraph after developing more nuanced, compelling connections and reasoning in subsequent body paragraphs. The insight that writers think in dynamic, non-linear ways has informed an entire body of scholarship focused on writers’ “self-regulation” habits, which illuminate a wide variety of ways that writers pick and choose which writing and revision strategies they should focus on at various points in their drafting process (Graham & Harris, 1994; Panadero, 2017; Zimmerman & Kitsantis, 1999).

GenAI writing tools, on the other hand, are not by themselves recursive, and are functionally far more linear than human writers in their composition processes. The vast majority of large language model (LLM) tools generate words, phrases, and sentences using “next-word” prediction, based upon the specific order of words from the human user’s prompt. In essence, the “writing goal” of the LLM is a static one: to predict the *most likely* series of words that should appear in its output, given the statistical relationships between words and phrases in its training dataset. This process cannot be described as anything other than a linear set of mathematical steps to determine what a user *probably* wants to see, based on their input prompt. As such, GenAI writing software cannot be said to “think” in a recursive manner—there are no hierarchies of task-representation to be found in its generative computation, and if there were, even the most advanced systems lack the ability to reflect upon, reorder, and transform these tasks. In fact, we might more accurately say that the quality of an LLM’s output is based primarily upon the *human* user’s ability to represent the context, audience, genre, and other constraints on writing style in their prompt, and to revise these representations as they continue to develop their ideas. In other words, it is still the human writer who is doing the complex thought-work that precedes the generation of text, not the language model.

SOPHISTICATED WRITING REQUIRES METACOGNITIVE REFLECTION

Following a similar line of inquiry, writing researchers have consistently found that the most sophisticated writers often draw upon a common observable thinking tool: metacognitive reflection. In essence, learners can most flexibly utilize their skills when they are able to think critically *about* their own thought processes and make informed decisions about how to regulate them while practicing new skills (Kuhn & Dean, 2004). As Rebecca S. Nowacek (2011) contends, writers who cultivate the metacognitive ability to name what they know—to identify particular strategies they use while writing, and articulate a rationale to justify *why* they are using them—are far more likely to transfer their knowledge of writing in one situation, task, or genre into others. For instance, students learning a new written genre in a classroom context—a book report, an analytic essay, a cover letter, etc.—will often by default draw on their knowledge and experiences of other writing tasks they have learned in previous classes or for other assignments. Gwen Gorzelsky et al. (2016) point out that students who show metacognitive self-awareness of what they are doing are able to more flexibly adapt to new writing tasks because they exercise a conscious conceptual control over their ideas and actions. This might, for instance, take the form of

applying writing strategies that they have previously used, comparing the new task to previous ones they have completed, and deliberately evaluating the usefulness of a writing practice as (or after) they use it.

While some GenAI chatbots have attempted to integrate more sophisticated algorithmic techniques that emulate metacognition—such as Anthropic AI’s Claude, which runs two or more “adversarial” LLMs against one another to assess the appropriateness of one’s output (Kundu et al., 2023), or recent developments in so-called “reasoning” models that utilize “chain-of-thought” techniques (Wei et al., 2022)—these scarcely resemble the diverse repertoire of metacognitive strategies identified by writing researchers. Instead, these more advanced GenAI models “simulate” a series of logical “reasoning” steps that are designed to textually resemble a spoken-aloud process of thought (Edwards, 2025). Even with this development, researchers at Apple have recently uncovered that the superficiality of this “reasoning” process continues to cause issues for even the most advanced models when they are faced with problems that require “outside-the-box” thinking (Shojaee et al., 2025). In sum, the immense difficulty of identifying a GenAI tool’s discrete, self-conscious control over its “decision-making” process, even by the most conservative estimation, is evidence that it cannot reason metacognitively about its text output in the same way that human writers can. Instead, we can point to the obvious fact that *human* users can (and do) consciously evaluate and reflect upon the output of an LLM, and use that understanding to modify their initial prompt or iterate it to produce something different – another instance of human writers doing the thinking that produces a language model’s response.

CONCLUSION

Beyond the arguments laid out above, it is important to remember that because LLM chatbots rely on word probability distributions to generate their output, their “writing” is, by definition, average. Based on the computational architecture of actually existing “AI” technologies, our definition of “thinking” would thus need to be refigured as a statistical averaging-out of a partial representation of human language-use on the internet. Without mincing words: this is an enormously impoverished definition of thinking—one which teachers, students, and practitioners of writing should not be at all satisfied or forced to accept. Instead, we must always keep in mind that human writers are still doing the thought-work that goes into producing text, even if it is automatically generated. Whether we want future generations of writers to rely upon large language models or not, we must always reinforce the idea that it is the person in front of the screen, not the hardware behind it, that is consciously making the complex cognitive decisions that inform how a piece of writing looks and sounds.

REFERENCES

- Baria, A. T., & Cross, K. (2021). *The brain is a computer is a brain: Neuroscience's internal debate and the social significance of the computational metaphor (preprint)*. *arXiv*. <https://arxiv.org/abs/2107.14042>
- Edwards, B. (2025, June 11). With the launch of o3-pro, let's talk about what AI "reasoning" actually does. *Ars Technica*. <https://arstechnica.com/ai/2025/06/with-the-launch-of-o3-pro-lets-talk-about-what-ai-reasoning-actually-does/>
- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365–387. <https://doi.org/10.2307/356600>
- Gorzelsky, G., Driscoll, D. L., Paszek, J., Jones, E., & Hayes, C. (2016). Chapter 8. Cultivating constructive metacognition: A new taxonomy for writing studies. In C. M. Anson & J. L. Moore (Eds.), *Critical transitions: Writing and the question of transfer* (pp. 215–246). The WAC Clearinghouse; University Press of Colorado. <https://doi.org/10.37514/PER-B.2016.0797.2.08>
- Graham, S., & Harris, K. R., (1994). *Self-regulation of learning and performance*. Routledge.
- Kelly, S. M. (2023, January 26). *ChatGPT passes exams from law and business schools*. CNN Business. <https://www.cnn.com/2023/01/26/tech/chatgpt-passes-exams/>
- Kuhn, D., & Dean, D. (2004). Metacognition: A bridge between cognitive psychology and educational practice. *Theory into Practice*, 43(4), 268–273.
- Kundu, S., Bai, Y., Kadavath, S., Askell, A., Callahan, A., Chen, A., Goldie, A., Balwit, A., Mirhoseini, A., McLean, B., Olsson, C., Evraets, C., Tran-Johnson, E., Durmus, E., Perez, E., Kernion, J., Kerr, J., Ndousse, K., Nguyen, K., ... Kaplan, J. (2023). *Specific versus general principles for constitutional AI (preprint)*. *arXiv*. <https://doi.org/10.48550/arXiv.2310.13798>
- Larsen, L. (2024, June 20). *GPT-5 will have "Ph.D.-level" intelligence*. *Digitaltrends*. <https://www.digitaltrends.com/computing/openai-says-gpt-5-will-be-phd-level/>
- Nowacek, R. S. (2011). *Agents of integration: Understanding transfer as a rhetorical act*. SIU Press.
- Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.00422>
- Shojaee, P., Mirzadeh, I., Alizadeh, K., Horton, M., Bengio, S., & Farajtabar, M. (2025). *The illusion of thinking: Understanding the strengths and limitations of reasoning models via the lens of problem complexity*. Apple. <https://ml-site.cdn-apple.com/papers/the-illusion-of-thinking.pdf>
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., Le, Q. V., & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, 35, 24824–24837.
- Zimmerman, B. J., & Kitsantas, A. (1999). Acquiring writing revision skill: Shifting from process to outcome self-regulatory goals. *Journal of Educational Psychology*, 91(2), 241–250. <https://doi.org/10.1037/0022-0663.91.2.241>