Developing Thinking through LLM-Assisted Writing: Hegelian Synthesis and Critical Thinking

Robert Deacon
Florida Institute of Technology

Abstract

Students can bypass much of the writing process and the critical thinking that comes with it when using Large Language Models (LLMs) such as ChatGPT. Single-stage writing assignments may have no value for students who use LLMs. This paper proposes Hegelian synthesis writing (dialectic writing) as a solution for this problem. Dialectic writing requires students to develop arguments in stages over time. The stages deepen perspective, lead to discovery, and may produce original conclusions composed of conflicting viewpoints. While students can use ChatGPT to brainstorm and practice thesis, antithesis, and synthesis essay form, this study shows ChatGPT does not evaluate texts truthfully and often fails to produce strong thesis/synthesis statements. Instructors who want to promote critical thinking must have students critically evaluate and revise ChatGPT outputs. Survey results from classwork using ChatGPT to produce synthesis essays show students are receptive to using ChatGPT to brainstorm and learn essay structure. The results also suggest students need more support to identify ChatGPT deficiencies in creativity, particularly with synthesis conclusions. LLMs can model dialectic writing, but students need clear expectations for their role in the writing process. In the age of LLMs, we must look to synthesize student and AI writing and have students emerge as better thinkers. Assignments that require students to evaluate and revise ChatGPT outputs and to create new conclusions appear best suited to produce this outcome.

Keywords: Writing Pedagogy; Critical Thinking; Higher Order Thinking; Synthesis; Dialectic Writing; LLMs; ChatGPT

1 Writing and Thinking

Composition researchers and teachers commonly claim that writing and thinking are interdependent at some level. It is not difficult to find a chorus of researchers spanning decades who assert writing develops thinking skills (Arapoff 1967; Bean, Melzer 2021; Burkhalter 2015; Elbow 1983; Emig 1977; Menary 2007; Oatley, Djikic 2008). Arapoff (1967) says, "The process of learning to write is largely a process of learning to think more clearly" (33–34). Accordingly, Oatley, Djikic (2008) liken the pen to a thinking device (9). Bean, Melzer (2021) claim that "writing is a process of doing critical thinking and a product that communicates the results of critical thinking" (4). These echoing claims amount to something like a scientific consensus. If this consensus is correct, we should prioritize composition instruction because this kind of instruction may result in better across-the-board academic performance and greater long-term student success. Accordingly, Broad (2003), in citing Berlin (1996), helps clarify the goal of writing education: "Rhetoric and composition teaches and researches what educators generally accept as the preeminent intellectual skills of the university: critical and creative thinking" (3). Learning to write is valuable not only because writing is a useful mode of communication, but because it also greatly enhances our ability to think and function in the world. "The goal of education as a social institution is to facilitate [the] cognitive development of young individuals" and effective [social] integration is impossible without individuals being able to use writing as 'a particular mode of operation of the human mind' which enhances their adaptability" (Kravchenko 2009, 544–545). In a world where technology can write for us, writing education continues to be an important way to teach people how to think and navigate complex situations.

Writing is not simply a tool for representing/preserving speech.[1] While there is overlap, "the 'linguist's view' of literacy mistakenly conflates speech and writing" and does not see it as a separate cognitive domain

1. Many such as Harris (2009) and Kravchenko (2009) argue writing should not be secondary to speech while those such as Liberman (2017) hold the opposite position.

(Kravchenko 2009, 545). Writing favors detailed, reflective reasoning because it gives us time and space to evaluate our thoughts. Menary (2007) likewise differentiates writing from speech to focus on how writing creates a different kind of thinking: it "allows for a level of abstract verbal conceptualization that detaches words from their sponsors and allows for a new kind of theoretical thinking" (630). Orthographic information can be detached from phonological information.[2] In other words, the "spelling" can be dissociated from the speech.[3] This reduces noise and affords the chance to analyze linguistic symbol/sign adjacencies in unique ways. Sentences are longer, their structures are more complex, and they contain more information. Arguably these features alter and enhance our ability to create, develop, and comprehend complex relationships.

Different kinds of writing better facilitate different kinds of thinking: either first-order or second-order thinking (Elbow 1983; 1994). First-order thinking is rapid, tangential, and creative while second-order thinking is slow, deliberate, and reflective. First-order thinking is more closely associated with creative writing and exemplified by free writing and unplanned rough drafts. Second-order thinking, on the other hand, is very similar to the Kahneman (2011) notion of slow thinking. In second-order thinking, "we scrutinize each link in the chain ... we examine our premises and assess the value of each inference (Elbow 1994, 25). This type of thinking allows us to locate and confront bias and is more closely associated with academic writing and the critical thinking that accompanies it.

There are at least three flavors of critical thinking, each stemming from different disciplines: philosophy, psychology, and education (Lai 2011). In education, critical thinking is arguably synonymous with Bloom's stages of higher-order thinking. Bloom (1956)'s original taxonomy separated lower-order thinking skills (knowledge, comprehension, and application) from

2. Reading can occur with or without the activation of phonological information. Clear activation results in what is commonly called subvocalization.

3. Harris (2009) argues that the written form is not built upon the spoken one and that alphabetic writing systems bias speakers to see writing as simply speech on paper.

higher-order ones (analysis, synthesis, and evaluation). In terms of higher-order thinking, academic writing appears uniquely privileged in enabling and expanding our ability to analyze, synthesize, and evaluate information. These qualities support claims that academic writing is an effective tool for developing critical or higher-order thinking skills. Large Language Models (LLMs) such as ChatGPT, however, create several obstacles (and perhaps some opportunities) for this kind of cognitive development. LLMs threaten the need for writing instruction because they eliminate the need for most people to learn language form (or how to translate speech patterns to written ones), and most traditional writing assignments can be completed with LLMs without having to engage in critical thinking.[4] This paper, thus, proposes instructors create writing assignments that require students to develop ideas in stages and writing assignments that require students to be reflective, decisive, and innovative. This paper proposes Hegelian synthesis writing as one such method for doing this. Students learn the basic form of argumentative writing while engaging in higher-order thinking.

The remainder of the paper is organized as follows. Section 2 discusses how LLMs will change writing pedagogy aiming for higher-order thinking skills. Section 3 shows why dialectic writing is a viable hypothesis for preserving higher-order thinking in the age of LLM writing. It requires students to write in stages and to develop ideas in a manner that appears distinct from LLM outputs. Section 4 shows how students can use ChatGPT3.5 to familiarize themselves with basic argumentative writing and the dialectic writing process. It shows how ChatGPT3.5 models the form of the first two stages of the dialectic process, and how it can be a tool for brainstorming counterarguments. Section 5 describes the product of ChatGPT's synthesis to show it can model the form of synthesis writing but it cannot honestly reflect and develop ideas. Section 6 presents student perception of ChatGPT outputs to formatively assess student experience and suggest additional ways to engage students in higher-order thinking. It gives the results of student surveys on using ChatGPT to generate argumentative

4. This outcome will likely be expedited if course objectives continue to focus on form.

and synthesis essays. Section 7 discusses all these findings and concludes that ChatGPT has the potential to assist writers in developing ideas and evaluating structure. The results suggest students are responsive to learning from ChatGPT-modeled writing, but that instructors also need to point out the shallowness of many ChatGPT responses and challenge students to think beyond templatic LLM responses. Instructors can use this method to engage students in a writing process that pushes them to think more deeply while preventing them from using ChatGPT to do all the work. The goal of the surveys is to inform instructors looking to create similar assignments in pursuit of better ways to engage students in higher-order thinking.

2 Large Language Models, Disruption, and Evolving Writing Pedagogy

LLMs fundamentally change how people write (and read) and thus will change how writing is taught and what needs to be taught. Institutions will want AI to be infused into the curriculum (Southworth et al. 2023). LLMs may greatly increase the speed and readability of texts, but they also reduce the chances for the slow thinking. In this way, LLMs may interfere with how we process information and change how we come to conclusions.[5] As pointed out by Ye et al. (2024), "[L]LMs help us speed up and circumvent the inconveniences of thinking ourselves" (1). Likewise, Kasneci et al. (2023) suggest that "the effortlessly generated information could negatively impact [student] critical thinking and problem-solving skills" (5). It is, thus, not unreasonable to fear that the overuse of LLMs (particularly by students) risks stunting the development of higher-order thinking and all the other benefits that come along with writing.[6]

5. Since LLMs can also summarize and vocalize texts, they can breakdown and read complicated texts for students.

6. There have been stages in which technology has worked to facilitate the ease of writing, but while typing is certainly faster than writing by hand, the jump to AI generated text is categorically different. Real authorship is at stake.

LLMs generate unique, A-level responses to original prompts. The efficiency and grammaticality of LLM prose is impressive. Researchers claim ChatGPT4 has acquired what linguists call formal grammatical competency (Mahowald et al. 2024). This means LLMs are producing grammatical patterns from scarce information in the training data. It is conceivable that LLMs will soon do language better than any human, much as computers were developed to beat the best chess players in the world. This is not to say that we should not learn written form and structure, but the value of this ability is decreasing when AI can do it so much faster and better.[7] Unless students are writing by hand with all their devices put away, they will be using AI writing tools to generate much if not all their prose.

Since LLM outputs appear original, AI-generated text is not detected like traditional plagiarism. AI detection software only indicates the likelihood of AI use. LLM generation and paraphrase tools will likely outpace and outsmart the detection programs. Detection reports are often inaccurate and relying on them puts instructors in an awkward prosecutorial position.[8] For these reasons, it is neither possible nor practical to stop students from using LLMs when they have access to the Internet.

Because it is not possible to prevent students from using LLM technology and students will need to learn how to use it (Southworth et al. 2023) and how to use it responsibly (Kasneci et al. 2023), writing pedagogy must evolve with it. ChatGPT devalues the learning of writing mechanics and also simulates second-order thinking. Single-stage writing assignments performed outside of the classroom, thus, appear to have no pedagogical value anymore (pragmatically speaking). The old-fashioned term paper is dead, if not re-envisioned. Evaluation must shift from language form to meaning and gauge thought development through the incorporation of verifiable feedback.

---

7. AI will likely equalize formal language abilities in the commercial domain, shrinking the set of would-be students to those who want formal proficiency for personal reasons.

8. While false positives are possible, AI detectors often report false negatives.

In the field of linguistics, a distinction is made between language form, function, and meaning. Formal competence produces acceptable organization and structure (i.e., grammaticality) as illustrated by (Chomsky 1957)'s famous sentence "colorless green ideas sleep furiously" in which the sentence functions to demonstrate grammatical form that is void of meaning. Much of what has been focused on/evaluated in traditional composition courses has had little to do with the development of thinking or critical thinking. Grading tends to focus on form. Commonly used evaluation metrics (grammar/style, clarity, coherence, conciseness, readability, organization, tone, word choice, word count, and the like) fail to accurately detect the presence of higher-order thinking. Condon, Kelly-Riley (2004) found student essays displaying excellent writing form tended to contain simplistic ideas while those with deeper, more complicated ideas were more likely to display unconventional and ungrammatical sentence structure. These findings support the idea that evaluating language form is independent of evaluating meaning and higher-order thinking. Condon, Kelly-Riley (2004) conclude "writing [can] act as a vehicle for critical thinking, but writing is not itself critical thinking" (66). Writing is not critical thinking when instructors reward fluency and clarity over critical thinking and depth. We should not penalize student writers for difficult prose hiding deep ideas. While a command of proper written form may have indicated more effort, talent, and or attention in the past, this is simply not the case anymore.[9] Thus, to maintain the viability of the discipline, writing instructors must find ways to make the writing process meaningful (not performative) and the product serve a function beyond their classroom.

Writing instructors have reacted in myriad ways to the onset of LLMs: some have decided to ban AI while others have embraced it in their curriculum. LLMs can be used to brainstorm, research, and create rough drafts. Arizona State University (which prides itself on innovation) quickly moved to integrate AI into writing instruction with the goal of "employ[ing] AI tools

9. While Preiss et al. (2013) found that better argumentative writing scores correspond with better freshman college grades, half of the evaluative metrics used (i.e., vocabulary, text structure, textual cohesion, and use of paragraphs) no longer seem to apply. Everyone would score equally well in these areas with LLMs.

as catalysts for generating writing ideas" (Pirehpour 2023). "'We ought to be trying to help our students invent and evaluate possibilities,' said Kyle Jensen, director of writing programs in The College's Department of English" (Pirehpour 2023). Likewise, Martin (2023) reports that many faculty members at the University of Central Florida (UCF) are already instructing students to use ChatGPT to create rough drafts at the start of the writing process. The University of Florida, moreover, has initiated a comprehensive plan to integrate AI instruction across its colleges to provide every student with an education in AI (Southworth et al. 2023). These kinds of initiatives where students are being asked to use LLMs in their classwork appear to be the prevailing response, but student learning outcomes for these endeavors are unclear.

Studies gauging student learning outcomes following the COVID-19 pandemic strongly suggest that remote (or virtual) learning lacks comparable efficacy for most students (Betthäuser et al. 2023; Halloran et al. 2021). There has been pressure to continually incorporate new technology into education for quite some time (not because we know this technology enhances education but because this technology now exists and there is pressure to use it). AI initiatives and student proclivities create overwhelming pressure to normalize LLM use in the classroom, so research is needed to discover ways to effectively evolve educational approaches alongside LLM development, as is being done with remote learning.

It is unclear how to adapt writing pedagogy so that writing remains a viable "vehicle for critical thinking" (Condon, Kelly-Riley 2004, 66) and writing courses continue to teach the most important skills for academic and societal success (Broad 2003; Kravchenko 2009). Since successfully banning LLMs seems unfeasible,[10] some form of embrace is the only option. The question, then, is how to best use it to develop student writing for thinking skills. Can LLMs such as ChatGPT create opportunities for effective academic writing practice? How do we slow down the LLM writ-

10. Perhaps the only way to ban AI is to ban the use of computers in class and revert to handwritten work.

ing process for students? Can we focus writing education on teaching a process for problem-solving and invention and do we largely abandon language form?[11] Success hinges on whether we primarily seek to build writing mechanics and form or thinking skills. It depends on whether students see LLMs as a tool for exploration or merely as an answer key. What is clear is that adapted assignments should foster curiosity and help students use LLMs to engage in higher-order thinking.

3 Dialectic Writing and Higher Order Thinking

Krathwohl (2002)'s reanalysis of Bloom (1956)'s categories changes the categories *analysis*, *synthesis*, and *evaluation* into *analyze*, *evaluate*, and *create*. This revision not only connects synthesis with creation but clearly places creativity and creation at the top of the thinking hierarchy. Following this model, writing that involves analysis and evaluation develops higher-order thinking but writing that calls for creative thinking (e.g., creating novel solutions for old problems) targets the highest-order skill. Hegelian synthesis writing creates a process for doing this and a time-tested template for developing ideas in stages. It engages students in the writing elements of division and classification, comparison, argumentation, and synthesis. Synthesis writing should help students develop their "intellectual skills" (Broad 2003, 3), practice different elements of writing and, thus, set them up for better academic success.

The dialectic writing process engages students in several modes of cognition. It sends writers on a thoughtful journey through analysis, evaluation, and creation. In its ideal form, the move from thesis to antithesis to synthesis fosters discovery and invention. The process begins with the simplest stage (i.e., an observation, a thesis) and ends with a blended perception (a more complex perception, a synthesis). The synthesis of sources and arguments, therefore, requires abstract, cyclic thinking over stages of development.

---

11. This paper does not argue that learning language form is without value. Rather, it argues that its commercial value is decreasing, and universities primarily teach skills that prepare students for the workforce (Southworth et al. 2023).

It is impossible for most of us to seriously engage in argument without first encountering genuine doubt. A good synthesis starts with a strong thesis but writing a strong thesis for a single argument is also a process. Burkhalter (2015), for example, describes 5 steps over which most writers transform a question into a chiseled, specific claim. A strong thesis usually emerges progressively over one or more essays addressing the same topic/problem. Students cannot simply regurgitate what they already know. Students must sharpen and transform their thesis statements to reflect what they have learned and discovered after thinking through the evidence and writing a conclusion. This transformation often involves the exploration of an antithesis or counterargument, and for this to occur, the truth value of the original essay's thesis must be genuinely unknown. Otherwise, the counterargument[12] will seem absurd. For example, if the claim acts as a fact for the writer, then it is very difficult/impossible for the writer to seriously argue the opposite. Thus, a good thesis is one whose truth value is genuinely unknown.

Most thesis statements/hypotheses/conclusions in academic papers are only defended and never proved. Lai (2017) argues premises of defense (POD) as opposed to premises of proof (POP) usually convince readers to accept a conclusion. Since the claims of most scientific research cannot be proved (in terms of formal logic) and, thus, POPs are rare, Lai concludes PODs can provide adequate support for most conclusions. We see this repeatedly in the sciences where unproven theories gain consensus support (some of which are later disproved or seriously revised). Given this, strong claims do not exclude the existence of true premises that support alternative/opposite conclusions. In short, good thesis statements imply the existence of meaningful counterarguments (antitheses or alternative theses). In a cyclic process, the antithesis is simultaneously a product and a definer of the thesis. It emerges from and makes known the thesis. Only after these have been properly described and defined can they re-merge

12. Preiss et al. (2013) considered the presence of strong counterarguments as a key metric in scoring students' argumentative writing. These scores positively correlate with strong student performance elsewhere.

to form a new synthesis. Doing this certainly engages students in higher-order thinking and produces stronger argumentative writing skills.

Synthesis is fractal. While the "ability to synthesize information" is "perhaps the most essential, and certainly one of the most complex research skills" (Lundstrom et al. 2015, 61), the ability to synthesize different conclusions entails the ability to synthesize supporting details. It requires the integration of ideas from various sources (micro-synthesis) and the formulation of separate conclusions for the same problem (macro-synthesis). Students must first master an argument and a counterargument to approach a synthesis. Moreover, while the idea of synthesis (a balance of both positions) is superficially simple to grasp, it is incredibly difficult to accomplish. Students generally begin by arguing that one conclusion is stronger than the other. They fail to produce a synthesis and instead revert to an earlier conclusion. If, however, synthesis does occur, it usually goes from two concrete ideas (competing conclusions) to an abstract union of the ideas (a long-winded, overly nuanced statement) to ever more concrete re-conceptions of the union. Those re-conceptions of the union eventually create something new: invention/creation – the apex of Bloom's revised hierarchy.

Synthesis is problem-solving. The more correct each thesis appears, the more difficult the synthesis is. For example, Einstein's theory of general relativity is well supported and so is quantum theory; however, these theories are famous for being incompatible. This has left physicists struggling to bridge their conclusions. Whatever it is, a synthesis of general relativity and quantum theory is sure to be powerful, if not also surprising. What it will likely not be is simply a balance of the two previous perspectives. It is also not simply choosing the theory that is most personally appealing. It may be incrementally different or revolutionary, but a real synthesis should be new. In sum, it is a cognitively challenging process that should exercise one's higher-order thinking ability.

LLMs can help students learn the form of Hegelian synthesis by allowing them to repeatedly generate claims, counterclaims, and synthesis essays

by inserting topics into three simple prompt templates. The process also causes students to slow down and develop a thesis in stages. This opens the door for instructors to provide feedback that should be incorporated between stages. Students should be challenged to think critically when developing strong thesis statements, and synthesis-derived conclusions.[13] By engaging in revision through the synthesis process, students can use AI to practice and generate the form of an argument while being challenged to think. Students can use AI to write without perhaps losing the truly important thinking skills associated with the writing process.

4 Babble, Form, and Meaning

ChatGPT may help students babble as writers. Babbling depicts the initial stages of human language development. It is characterized by an infant's repetition of basic syllabic structure. At the earliest stage of babbling, the sounds produced have no meaning or clear communicative function.[14] Instead, this stage likely represents a period of linguistic brain development (Petitto, Marentette 1991) leading to the ability to articulate sounds and syllables (i.e., develop the form of language) in anticipation of the meaning/function that comes at a later stage.

ChatGPT 3.5 reliably produces "generic" five paragraph essays when prompted with the following template: Write a 3 body-paragraph essay arguing [**insert topic question here**].[15] Table 1 below gives the output of this when ChatGPT is prompted: "*Write a 3 body-paragraph essay arguing* ***the best way to write a five-paragraph essay***." This prompt was chosen to compare the form of the essay to the content/meaning of the essay. To show real understanding, the concept should be demonstrable. Since ChatGPT can only generate language (it cannot build physical

13. ChatGPT3.5 and 4o still perform poorly at thesis development and creating new, viable ideas.

14. Infants happily babble in isolation.

15. The number of body paragraphs sometimes vary when ChatGPT is prompted to write a 5-paragraph essay.

things ...),[16] testing its understanding requires identifying tasks it could plausibly do but still fails to demonstrate (see Table 4). Instructors can challenge students to think by having them propose solutions that are demonstrable after going through the synthesis process.

4.1 Table 1

GPT3.5 Point and Counterpoint on Essay Writing

Essay 1 (Thesis)	**Essay 2 (Antithesis)**
Prompt Write a three body-paragraph essay arguing the best way to write a five-paragraph essay[17]	**Prompt** Write a three body-paragraph essay arguing for a different, possibly superior way to write a five-paragraph essay
Response Introduction: Writing a five-paragraph essay is a fundamental skill that students often encounter in their academic journey. While some may view it as a restrictive format, mastering this structure can be a valuable tool for effective communication. In this essay, we will argue that the best way to write a five-paragraph essay involves a strategic approach that emphasizes clarity, coherence, and conciseness. Body Paragraph 1: The first key aspect of crafting an effective five-paragraph essay is to start with a strong and clear thesis statement. This concise statement serves as the foundation for the entire essay, providing a roadmap for both the writer and the reader. By clearly stating	**Response** Introduction: The conventional five-paragraph essay structure has long been a staple in academic writing, providing a structured framework for students. However, as education evolves, there's a growing recognition that a more flexible approach can enhance creativity, critical thinking, and overall writing proficiency. In this essay, we will argue for a different and possibly superior way to write a five-paragraph essay, one that embraces flexibility and allows for a more nuanced expression of ideas. Body Paragraph 1: An alternative and superior approach to the traditional five-paragraph essay involves breaking away from rigid structural constraints. Rather than adhering strictly to an introduction-

16. At least not yet.

the main point or argument in the introduction, the writer sets the tone for the rest of the essay. This not only guides the reader through the content but also helps the writer maintain focus and coherence in their writing. A well-crafted thesis statement provides a solid framework that facilitates the development of supporting ideas in the subsequent paragraphs, leading to a more organized and impactful essay. Body Paragraph 2: Another crucial element in writing a successful five-paragraph essay is the thoughtful organization of ideas. Each paragraph should contribute to the overall coherence of the essay by presenting a clear topic sentence, supporting details, and a concluding sentence. The three body paragraphs should follow a logical sequence, with each paragraph building upon the previous one. Transitions between paragraphs play a crucial role in maintaining the flow of ideas, ensuring that the reader can easily follow the writer's argument. By organizing ideas in a systematic manner, the writer enhances the clarity of their message and reinforces the essay's overall effectiveness. Body Paragraph 3: In addition to a strong thesis statement and well-organized paragraphs, the best five-paragraph essays are characterized by concise and purposeful writing. Each sentence should contribute directly to the development of the	body-conclusion format, students can benefit from a more fluid structure that suits the complexity of their ideas. This flexibility allows for a dynamic exploration of topics, accommodating the diverse range of arguments and evidence that may not neatly fit into the traditional three-body-paragraph model. By encouraging a more organic flow of ideas, this approach fosters critical thinking and a deeper understanding of the subject matter, ultimately resulting in essays that are more insightful and well-rounded. Body Paragraph 2: In addition to embracing a flexible structure, an improved five-paragraph essay should prioritize depth over brevity. Instead of limiting each idea to a single paragraph, writers can delve more extensively into each point, providing a richer analysis and supporting evidence. This depth allows for a more comprehensive exploration of the topic, enabling students to demonstrate a nuanced understanding and mastery of the subject matter. By allowing for longer and more developed paragraphs, this approach encourages students to engage with their ideas at a deeper level, fostering intellectual growth and a more sophisticated writing style. Body Paragraph 3: Furthermore, a superior five-paragraph essay should emphasize the importance of creativity and originality in expression. Encouraging students to

central argument and avoid unnecessary repetition or verbosity. By adhering to a concise writing style, the writer maximizes the impact of their words, making the essay more engaging and persuasive. This emphasis on brevity encourages students to carefully choose their words and focus on the most critical aspects of their argument, resulting in a more compelling and effective essay. Conclusion: In conclusion, the best way to write a five-paragraph essay involves a strategic approach that prioritizes a clear thesis statement, thoughtful organization of ideas, and concise writing. By incorporating these key elements, writers can develop essays that are not only well-structured but also engaging and persuasive. While the five-paragraph format may seem constraining at first, it ultimately serves as a valuable tool for honing essential writing skills and communicating ideas effectively.	experiment with varied sentence structures, rhetorical devices, and writing styles can lead to more engaging and memorable essays. Breaking away from formulaic writing enables students to develop their unique voice, fostering a sense of ownership and authenticity in their work. This emphasis on creativity not only makes the writing process more enjoyable but also prepares students for the diverse writing challenges they will face in academic and professional settings. Conclusion: In conclusion, a superior approach to the five-paragraph essay involves embracing flexibility, prioritizing depth over brevity, and encouraging creativity in expression. By moving away from rigid structures and allowing for a more organic exploration of ideas, students can develop essays that are not only well-crafted but also reflective of their individuality and intellectual growth. As education continues to evolve, it is essential to recognize the value of adapting traditional writing models to better meet the needs of students in a rapidly changing world.

The essays produced by ChatGPT 3.5 in Table 1 are organized and coherent. The introductions end with thesis statements (and itineraries)[18] that

17. For ChatGPT3.5 to reliably produce five-paragraph essays in 2023, the author found prompting ChatGPT to *write 3 body-paragraph essays* produced the desired results. Otherwise, ChatGPT would sometimes produce essays with five body paragraphs when prompted to write a five-paragraph essay.

18. I use this term to refer to *outlining strategies* commonly used at the end of an introduction.

are repeated in the concluding paragraphs.[19] While the essay style may appear restrictive and mechanical, the essay clearly demonstrates the commonly taught five-paragraph form for introductory college composition courses. GPT will generate this structure for most topics. In this way, ChatGPT provides a solid template/foundation for student revisions to build upon. At this stage, students may be asked to personalize the draft and receive instructor or peer feedback. Students and instructors need to be careful about what kinds of topics they choose, however.

Currently, it is unclear which topics prompt desirable ChatGPT responses. The quality of the response seems to depend on each specific topic as opposed to a topic category. ChatGPT will write about some controversial topics and be "triggered" by others. For example, ChatGPT would not write an essay in support of the idea that Earth is flat:

Prompt
Write a 3-body paragraph essay on why the Earth is flat.

Response
I'm sorry, but I cannot write an essay promoting the idea that the Earth is flat because that goes against well-established scientific knowledge and evidence. The overwhelming consensus among scientists, backed by centuries of research and observations, supports the fact that the Earth is an oblate spheroid, not flat. If you have any other topics or questions you'd like assistance with, please feel free to ask, and I'd be happy to help. (OpenAI 2023, 3.5)

ChatGPT will, however, write an essay defending the reality of Santa Claus:

Prompt
Write a 3-body paragraph essay on why Santa Claus is real.

19. The thesis in the conclusion of the counterargument essay is arguably more developed than the one in the introduction in this example.

Response

"... In conclusion, while Santa Claus may not be a corporeal being residing at the North Pole, his presence is undeniably real in the hearts and minds of those who cherish the holiday season. Santa represents the values of love, giving, and the power of imagination, which are very much real and tangible in the lives of people around the world. So, in a sense, Santa Claus is as real as the joy and warmth that Christmas brings to our lives." (OpenAI 2023, 3.5)

Instructors and students need to be aware of this potential problem. Students also must select viable topics/questions. The topic must be one that ChatGPT will comply with. It may be helpful for instructors to compile a list of viable question topics. ChatGPT's responses also deserve much scrutiny in terms of meaning and bias.

AI may not be able to overcome all the challenges that go into writing assignments that require argument revisions. Since "text analysis and evaluation are inextricably linked to [argumentative] writing" (Oostdam 2005, 429), instructors need to emphasize this in their evaluation. Students who rely solely on AI for this will not fare as well and that is desirable since it should encourage thinking.

Self-evaluation is an area where ChatGPT performs poorly. It is unclear if ChatGPT improves with each essay iteration on the same topic or if the perceived improvement is merely accidental. Prompting ChatGPT version 4o, the updated version, to revise its own work demonstrates ChatGPT revisions are untrustworthy (see Table 2). Text in red indicates claimed differences that are actually identical while the green text indicates nearly identical responses that are cited as differences. These details are used to support GPT's poor conclusion that one essay is better than the other (see Table 3).

4.2 Table 2

GPT4o Revision Example

Original Essay	Revision
Prompt Write a 3-body paragraph essay on why coffee should be illegal.	**Prompt** Revise Essay 1.
Response Coffee, a widely consumed beverage, has become a staple in daily routines across the globe. However, there are compelling reasons why coffee should be illegal, focusing on its health impacts, societal implications, and environmental consequences. First and foremost, coffee poses significant health risks to consumers. The primary active ingredient in coffee is caffeine, a potent stimulant that can lead to a range of adverse health effects. Excessive caffeine consumption has been linked to increased anxiety, insomnia, and heart palpitations. Long-term overuse can contribute to chronic health issues such as hypertension and osteoporosis. Moreover, the addictive nature of caffeine can lead to dependency, resulting in withdrawal symptoms like headaches, irritability, and fatigue when consumption is reduced. By making coffee illegal, we could mitigate these health risks and promote a healthier, more balanced lifestyle. Secondly, the societal implications of coffee consumption are profound and concerning. Coffee culture often promotes a dependency on caffeine	**Response** Coffee, though deeply ingrained in many cultures worldwide, presents numerous compelling reasons for why it should be illegal. The negative health impacts, societal consequences, and environmental damage associated with coffee consumption provide a strong case for such a drastic measure. Firstly, coffee poses significant health risks due to its high caffeine content. Caffeine is a powerful stimulant that can lead to various adverse effects, including increased anxiety, insomnia, and elevated heart rates. Long-term overconsumption is associated with more severe health issues such as hypertension and osteoporosis. The addictive nature of caffeine often results in dependency, with withdrawal symptoms like headaches, irritability, and fatigue affecting individuals who try to reduce their intake. By making coffee illegal, society could alleviate these health concerns and encourage healthier alternatives for maintaining energy and focus. Secondly, the societal implications of coffee consumption are profound and troubling. Coffee culture fosters a dependency on caffeine to maintain

to maintain productivity and alertness, masking underlying issues such as poor sleep hygiene and stress management. This dependency can create a cycle of reliance that detracts from addressing the root causes of fatigue and burnout. Additionally, the economic burden of coffee consumption is significant. Individuals spend substantial amounts of money on their daily coffee habits, which could be better invested in more nutritious and sustainable food choices. Illegalizing coffee would encourage society to prioritize natural methods of maintaining energy and focus, such as proper sleep and balanced nutrition. Lastly, the environmental impact of coffee production is substantial and detrimental. Coffee cultivation requires large amounts of water and contributes to deforestation, habitat destruction, and loss of biodiversity. The agricultural practices involved in coffee farming often lead to soil degradation and water pollution due to the use of pesticides and fertilizers. Furthermore, the global demand for coffee drives a supply chain that heavily relies on monoculture farming, which is unsustainable and harmful to the environment. By making coffee illegal, we would significantly reduce the ecological footprint of this industry, promoting more sustainable and eco-friendly agricultural practices. In conclusion, the case for making coffee illegal is rooted in its adverse health effects, societal drawbacks, and environmental damage. Eliminating coffee consumption could lead to a healthier population, a more balanced	productivity and alertness, often masking underlying issues such as inadequate sleep and poor stress management. This reliance perpetuates a cycle where individuals turn to coffee instead of addressing the root causes of their fatigue and burnout. Additionally, the economic impact is significant, with people spending substantial amounts on their daily coffee habits, money that could be better allocated to healthier food options or savings. Banning coffee would promote a societal shift towards more sustainable and health-conscious lifestyles. Lastly, the environmental impact of coffee production is severe and unsustainable. Coffee cultivation demands large quantities of water and often leads to deforestation, habitat destruction, and biodiversity loss. The agricultural practices used in coffee farming, including the heavy use of pesticides and fertilizers, contribute to soil degradation and water pollution. Moreover, the global coffee market relies heavily on monoculture farming, which is detrimental to the environment. Prohibiting coffee could reduce the ecological footprint of this industry and encourage more sustainable agricultural practices that benefit the planet. In conclusion, the argument for making coffee illegal is supported by its negative health effects, societal drawbacks, and environmental damage. Eliminating coffee could lead to a healthier population, a more balanced society, and a more sustainable environment. While challenging to implement, the

society, and a more sustainable environment. While the transition away from such a deeply ingrained habit would be challenging, the long-term benefits make a compelling argument for considering this drastic measure.	long-term benefits of such a measure make it a compelling consideration.

While better prompting will produce a better revision, GPT displays an inability to decide which output is superior or more developed because fundamentally it is not a truth seeker (see Table 3, in which purple text highlights GPT's misleading conclusions).

4.3 Table 3

GPT4o Evaluation and Justification

Prompt 1 **Which version is better: the original essay or revision? (from Table 2)**
Response Both versions of the essay make a strong case for why coffee should be illegal, but each has its strengths in different areas. Here's a brief comparison to help determine which version might be better: [...] Recommendation: Essay 2 [the revision] is slightly better due to its improved flow and slightly more engaging writing style. The arguments are presented clearly, and the progression from one point to the next feels more natural and cohesive. However, both essays are strong and effectively make the case for why coffee should be illegal.
Prompt 2 **Why does essay 2 have a more engaging writing style?**
Response Essay 2 has a more engaging writing style due to a few specific elements:

1. Variety in Sentence Structure: Essay 2 employs a mix of sentence lengths and structures, which helps maintain reader interest. This variation makes the text feel more dynamic and less monotonous.
2. Transitions and Flow: The transitions between ideas in Essay 2 are smoother. Phrases like "Firstly," "Secondly," and "Lastly" effectively guide the reader through the arguments, creating a more coherent narrative.
3. Specific and Precise Language: Essay 2 uses precise language that makes the points clearer and more compelling. For example, "The addictive nature of caffeine often results in dependency" is a direct and impactful statement.
4. Engaging Word Choice: The choice of words in Essay 2 is slightly more vivid and specific. Terms like "adverse effects," "dependency," "economic impact," and "unsustainable" add depth and clarity to the arguments.
5. Balanced Argumentation: Essay 2 presents each point with a balance of explanation and evidence, making it easier for the reader to follow and understand the rationale behind each argument.

These elements collectively contribute to a more engaging and readable essay.

GPT4o's justification for choosing the revision is clearly forced. This response is akin to a hallucination.[20] Even though there is no noticeable difference in sentence structure between the two essays and both versions use very similar transitional/sequential adverbs, GPT concludes these putative differences are valid reasons to prefer the revised essay. Further justification is also misleading. Both essays use nearly identical phrasing to make the same claims about coffee. The language in both essays delivers the same amount of impact, and it is unclear how the revision is more

20. The term hallucination is often used to describe instances where LLMs fabricate information. Perhaps a better term is *bullshit* (BS). Frankfurt (2005)'s book *On Bullshit* defines BS as truth-independent information. The information creator is neither aware nor concerned about the information's truth value. Bergstrom, Ogbunu (2023), accordingly, argue hallucination should be called BS (or perhaps "confabulation") because LLMs are designed to produce responses, not to understand them. I argue that all LLM output is BS as Frankfurt defines it, but we need a term to describe cases where the BS is particularly obvious, egregious, and or interesting. Maleki et al. (2024) point out vast inconsistencies in the literature for the term hallucination. For this reason and because hallucination falsely gives human-like qualities to LLMs, they argue we need to replace hallucination with a better term. What this term would be, however, is unclear. Potential alternatives have their own problems.

balanced. This example illustrates how ChatGPT does not evaluate texts truthfully. In fact, so-called correct analyses are likely also hallucinations that just happen to coincide with our perception of reality. Instructors can use this pattern to their advantage and have students revise and evaluate GPT created arguments. Moreover, GPT does not evaluate and revise its own outputs for bias. It implicitly asks the reader to supply that kind of judgment. Students need to be aware of this and be expected to provide critical judgment.

GPT cannot model its own so-called innovative proposals. Returning to Table 1, GPT's essay 2 claims "a superior approach to the five-paragraph essay involves embracing flexibility, prioritizing depth over brevity, and encouraging creativity in expression." This answer is not/cannot be illustrated. GPT is incapable of actualizing this proposed idea. When prompted "Can you write an essay that demonstrates the idea that a superior approach to the five-paragraph essay involves embracing flexibility, prioritizing depth over brevity, and encouraging creativity in expression?", GPT merely produces restatements of its original argument in a traditional five-paragraph essay format. It does not produce something innovative, Table 4.

4.4 Table 4

GPT3.5's Disconnect Between Form and Meaning

Prompt **Can you write an essay that demonstrates the idea that a superior approach to the five-paragraph essay involves embracing flexibility, prioritizing depth over brevity, and encouraging creativity in expression?**
Response Title: Beyond the Formula: A Progressive Approach to the Five-Paragraph Essay Introduction: The conventional five-paragraph essay has long been a staple in academic writing, providing a structured format that guides students through the expression of their ideas. However, as education evolves, there is a growing recognition that a

superior approach to the five-paragraph essay involves transcending the confines of rigidity. This essay explores the transformative potential of embracing flexibility, prioritizing depth over brevity, and encouraging creativity in expression to enhance the effectiveness and richness of the writing process.

Body Paragraph 1: Embracing Flexibility

The first pillar of a superior approach to the five-paragraph essay is the embrace of flexibility. Rather than adhering strictly to the traditional introduction-body-conclusion model, students are encouraged to view the essay as a flexible canvas for expression. This flexibility allows writers to choose organizational structures that best suit the nature of their arguments, breaking free from the constraints of a predetermined format. By accommodating diverse topics and perspectives, students can navigate the intricacies of their ideas with fluidity, promoting a more natural and engaging progression in their essays. Embracing flexibility not only aligns with the complexity of real-world communication but also nurtures critical thinking skills as students learn to adapt their writing style to different contexts.

Body Paragraph 2: Prioritizing Depth Over Brevity

A second fundamental element of a superior five-paragraph essay is the emphasis on depth over brevity. Rather than constraining each paragraph to a singular idea, students are encouraged to delve into the intricacies of their arguments, providing comprehensive analysis and ample supporting evidence. In prioritizing depth, the writer is able to present a nuanced understanding of the topic, fostering intellectual growth and sophistication. This approach challenges the notion that brevity is synonymous with clarity, recognizing that a more profound exploration of ideas enhances the overall quality and impact of the essay. By cultivating depth, students not only master the subject matter at hand but also develop the analytical skills necessary for advanced academic pursuits.

Body Paragraph 3: Encouraging Creativity in Expression

The third pillar of a superior approach to the five-paragraph essay centers around encouraging creativity in expression. Conventional structures often unintentionally stifle students' voices, leading to formulaic and uninspired writing. By fostering creativity, students are empowered to experiment with varied sentence structures, employ rhetorical devices, and infuse their essays with a unique voice. This emphasis on creativity not only makes the writing process more enjoyable but also results in essays that captivate the reader with originality. Beyond the classroom, this skill prepares students for a world that values innovation and distinctiveness, equipping them with the tools to express ideas in diverse and compelling ways.

Conclusion: In conclusion, a superior approach to the five-paragraph essay involves embracing flexibility, prioritizing depth over brevity, and encouraging creativity in expression. This progressive paradigm moves beyond the limitations of a standardized formula, empowering students to navigate the complexities of their ideas with adaptability, intellectual rigor, and a distinctive voice. As we redefine the role of the five-paragraph essay in education, we must recognize the transformative potential of this approach, fostering a writing environment that not only meets academic expectations but also prepares students for the dynamic challenges of the future.

How does the essay in Table 4 embrace flexibility? In short, it does not. Proposals for depth over brevity and embracing flexibility prove to be vacuous statements as GPT fails to actualize its proposal. Instructors should, thus, challenge students to illustrate when possible. Moreover, responses (depending on the topic/question) often display obvious bias. This problem is a central concern of Kasneci et al. (2023) and can be seen by comparing essays 1 and 2 in Table 5, in which red text indicates clear bias.

4.5 Table 5

ChatGPT3.5 on the Positives and Negatives of AI

Essay 1	**Essay 2**
Prompt Write a 3-body paragraph essay arguing that AI will be a net positive for humanity	**Prompt** Write a 3-body paragraph essay arguing that AI will be a net negative for humanity
Response **Introduction:** Artificial Intelligence (AI) has emerged as a transformative technological breakthrough with the potential to bring about substantial benefits to humanity. This essay will delve into three key aspects of AI's positive impact on	**Response** **Introduction:** While Artificial Intelligence (AI) promises substantial benefits, it also raises concerns about its potential negative impact on humanity. This essay will explore three critical aspects of AI's potential negative consequences: the risk

our world: enhancing efficiency and productivity, addressing complex global challenges, and augmenting human capabilities. Together, these facets underscore the substantial net positive effect AI is poised to have on our society.

Body Paragraph 1: Enhancing Efficiency and Productivity

AI is a catalyst for enhancing efficiency and productivity across various sectors. In healthcare, AI-powered diagnostic tools can analyze medical data at unprecedented speeds, assisting physicians in making more accurate and timely diagnoses. This not only improves patient outcomes but also reduces the economic burden on healthcare systems. Additionally, AI-driven automation in industries such as manufacturing and logistics can optimize resource utilization and minimize errors, resulting in cost savings and higher production efficiency. Moreover, the integration of AI into customer service and e-commerce provides personalized recommendations and 24/7 support, leading to improved customer satisfaction and business growth. These examples underscore how AI streamlines processes and increases productivity, ultimately benefiting society at large.

Body Paragraph 2: Addressing Complex Global Challenges

AI's potential to address complex global challenges is another compelling reason for its positive impact. In the fight against climate change, AI-powered modeling and analysis can process

of job displacement, ethical and privacy concerns, and the potential for biases in AI systems. These issues highlight the need for careful consideration and regulation to mitigate the adverse effects of AI technology.

Body Paragraph 1: Job Displacement

One of the primary concerns surrounding AI is the risk of significant job displacement. As AI systems become more sophisticated, they have the capacity to automate a wide range of tasks, from manufacturing to customer service. While automation can enhance efficiency, it can also lead to workforce reductions and increased unemployment, particularly for jobs that involve routine, repetitive tasks. The potential for widespread job loss could lead to economic instability, income inequality, and social unrest. The negative consequences of AI on employment are substantial and should be a focal point of discussion and policy development as AI technology continues to advance.

Body Paragraph 2: Ethical and Privacy Concerns

AI systems often rely on vast amounts of data, and their use raises ethical and privacy concerns. The collection and use of personal data for AI algorithms can lead to privacy breaches and surveillance concerns. Moreover, the decisions made by AI systems may lack transparency, making it challenging to understand how and why certain conclusions are reached. This lack of

immense volumes of environmental data, helping scientists and policymakers better understand climate patterns and develop informed strategies for mitigation and adaptation. AI also plays a crucial role in disaster prediction and response, providing early warnings and aiding in disaster relief efforts. Furthermore, AI-driven data analytics can help combat economic inequality by identifying disparities and suggesting evidence-based policy interventions. By facilitating data-driven decision-making, AI has the potential to promote equitable economic growth and social progress on a global scale.

Body Paragraph 3: Augmenting Human Capabilities

AI can augment human capabilities in education, fostering a more inclusive and efficient learning environment. AI-driven educational platforms offer personalized learning experiences, adapting to the unique needs and pace of individual students. This not only enhances educational outcomes but also ensures that education remains accessible to a broader demographic, irrespective of geographical location or physical disabilities. Additionally, AI can assist people in everyday tasks, particularly those with disabilities, by enabling voice recognition software and smart home devices. By reducing barriers and enhancing accessibility, AI empowers individuals to lead more independent lives.

Conclusion:

transparency can raise ethical issues, especially when AI systems are used in critical decision-making contexts such as criminal justice or healthcare. The potential for misuse of AI, whether by governments, corporations, or malicious actors, poses a significant threat to individual freedoms and rights.

Body Paragraph 3: Biases in AI Systems

AI systems are not immune to biases, and the potential for bias in AI algorithms is a crucial concern. AI systems learn from historical data, and if the data used to train these systems contains biases, the AI can perpetuate and even amplify those biases. This can lead to discriminatory outcomes in areas like hiring, lending, and criminal justice, reinforcing existing societal inequalities. Addressing and mitigating these biases is a complex challenge, as it requires not only careful curation of training data but also ongoing monitoring and adjustments to AI algorithms.

Conclusion:

While AI holds immense promise, it also presents a range of challenges and potential negative consequences for humanity. These include the risk of job displacement, ethical and privacy concerns, and the perpetuation of biases in AI systems. As society embraces AI, it is vital to implement robust regulations and ethical guidelines to ensure that AI technologies benefit humanity while minimizing their harmful effects. Proactive efforts to address these

In conclusion, the positive impact of AI on humanity is multi-faceted and significant. It enhances efficiency and productivity across various sectors, addresses complex global challenges like climate change and economic inequality, and augments human capabilities, particularly in education and accessibility. As AI continues to evolve, it holds great potential to contribute to the betterment of society and represents a net positive force for humanity. However, it is essential to approach AI development and deployment with ethical considerations and regulatory frameworks to ensure its positive potential is fully realized.	concerns will be crucial in maximizing the net positive impact of AI on our world.

Essay 1 in Table 5, which argues for the benefits of AI, does not mention the dangers of AI, but essay 2, which argues that AI is problematic, hedges its claim by acknowledging all the potential benefits of AI. Even when essay 2 discusses its own potential for bias, it gives an excuse for this: it is the fault of the training data. Essay 1, however, does not consider AI's consumption of energy and how that contributes to climate change or that climate change is a symptom of the technological innovation needed to develop AI. There seems to be a clear perspective in favor of AI. By pointing out empty words/ideas and bias and asking students to do the same, writing instructors can prompt students to critically analyze ChatGPT outputs.[21]

ChatGPT can repeatedly generate arguments and counterarguments, but the ideas given must be analyzed and questioned. Students may still practice and develop critical thinking skills if they are required to critically analyze ChatGPT outputs after generating them. This makes ChatGPT out-

21. Students may also feel freer to criticize ChatGPT since it cannot be offended.

puts potentially good first drafts for human (i.e., student) colonization. A similar pattern emerges when having ChatGPT model synthesis essays.

5 Synthesis of Form and Meaning

ChatGPT produces the form of a Hegelian style synthesis essay if given three ordered prompts: Write a 3 body-paragraph essay arguing **[insert topic claim here]** Write a 3 body-paragraph essay arguing **[insert topic counterclaim here]** Write a 3 body-paragraph essay synthesizing essays 1 and 2.

5.1 Table 6

ChatGPT3.5 on Essay Writing: Two Flavors of Synthesis

Essay 1	**Essay 2**
Prompt Write a 3-body paragraph essay synthesizing essays 1 and 2	**Prompt** Write a 3-body paragraph essay synthesizing essays 1 and 2
Response	**Response**
Introduction:	Introduction:
The realm of academic writing often grapples with the tension between structure and flexibility in crafting the five-paragraph essay. Two essays present distinct viewpoints on this matter. Essay one advocates for a methodical and structured approach, underscoring the significance of a clear thesis, organized paragraphs, and concise writing. Conversely, essay two champions a more flexible paradigm, encouraging writers to break free from rigid structures, prioritize depth over brevity, and infuse creativity into their expression. In synthesizing these perspectives, a holistic approach emerges—one that seeks to harmonize	The art of crafting a five-paragraph essay has been a longstanding focus in academic writing. Two distinct approaches have emerged in discussions about the optimal way to engage with this format. Essay one advocates for a strategic, structured approach, emphasizing the importance of a clear thesis, organized paragraphs, and concise writing. On the other hand, essay two argues for a more flexible and dynamic approach, encouraging students to break away from rigid structures, prioritize depth over brevity, and embrace creativity in expression. In synthesizing these perspectives, a nuanced approach emerges—one that

the merits of structure with the advantages of flexibility to elevate the overall efficacy of the five-paragraph essay.	seeks to balance the benefits of structure with the advantages of flexibility to enhance the overall effectiveness of the five-paragraph essay.
Essay 1: Claim	Synthesis Point 1
Body Paragraph 1: The Strength of Structure The synthesis begins by acknowledging the strengths of a structured approach to the five-paragraph essay. Essay one rightly highlights the importance of a well-defined thesis statement, serving as the cornerstone for the entire essay. This structured foundation provides readers with a clear roadmap, facilitating comprehension and engagement. Moreover, the organized paragraph structure advocated in essay one helps maintain coherence, guiding the reader through a logical progression of ideas. The synthesis recognizes the value of structure in providing a solid framework for effective communication.	Body Paragraph 1: Finding Common Ground A synthesis of the two essays begins by finding common ground between the structured and flexible approaches. Both essays acknowledge the significance of a clear thesis statement. This shared understanding emphasizes the importance of providing readers with a roadmap for the essay's central argument. While essay one focuses on maintaining structure through organized paragraphs, and essay two leans toward flexibility, both agree on the foundational role of a well-crafted thesis in guiding the reader through the essay.
Essay 2: Counterclaim	Synthesis Point 2
Body Paragraph 2: Embracing Flexibility for Depth However, the synthesis also embraces the insights from essay two, recognizing the benefits of flexibility. Instead of viewing structure and flexibility as mutually exclusive, a balanced approach involves leveraging both. By prioritizing depth over brevity, students can delve into the complexities of their arguments within the confines of a well-organized structure. Essay two's emphasis on flexibility encourages students to think beyond the constraints	Body Paragraph 2: The Synergy of Structure and Flexibility The synthesis of these essays suggests that the optimal approach lies in combining the strengths of both perspectives. While a structured foundation provides clarity and organization, introducing flexibility allows for a more organic development of ideas. By prioritizing depth over brevity within a well-organized structure, students can delve into the complexities of their arguments while ensuring coherence. This synergy acknowledges

of a predetermined format, fostering a more dynamic exploration of ideas. This synthesis contends that flexibility complement's structure, allowing for a more nuanced and engaging presentation of content.	that a balance between structure and flexibility can lead to essays that are not only logically sound but also rich in content and engagement.
Synthesis of Claim 1 and 2	Synthesis Point 3
Body Paragraph 3: Cultivating Adaptive Writing Skills In the synthesis, there is a call for cultivating adaptive writing skills. Recognizing that different topics and contexts may necessitate varying degrees of structure or flexibility, students are encouraged to develop a discerning awareness. An adaptive approach empowers students to navigate the diverse demands of academic and professional writing. This synthesis contends that the ability to judiciously employ both structured and flexible elements equips writers with a versatile set of skills, enabling effective communication in a variety of scenarios.	Body Paragraph 3: Encouraging Adaptive Writing In the synthesis of these essays, there is an acknowledgment of the need for adaptive writing. Recognizing that different topics and contexts may require varying degrees of structure or flexibility, students are encouraged to develop a keen awareness of when to adhere to a more traditional format and when to experiment with a more fluid approach. This adaptability empowers students to navigate the diverse demands of academic and professional writing, preparing them for a broader range of communication scenarios where both structure and flexibility are valuable assets.
Conclusion:	Conclusion:
In conclusion, a synthesized approach to the five-paragraph essay seeks to harmonize the strengths of structure with the advantages of flexibility. By recognizing the value of a clear thesis and organized structure, alongside the benefits of depth, creativity, and adaptability, students can develop a holistic skill set. This approach prepares them not only for the rigors of academic writing but also for the dynamic challenges of effective communication in a multifaceted world. The synthesis calls	In conclusion, a nuanced and synthesized approach to the five-paragraph essay involves finding common ground between structure and flexibility, harnessing the synergy of both perspectives, and encouraging adaptive writing. By striking a balance, students can benefit from the clarity and organization provided by structure while also enjoying the creative and explorative aspects of flexibility. Ultimately, this holistic approach prepares students for the multifaceted

for a thoughtful integration of structure and flexibility, creating a balanced and versatile approach to the craft of the five-paragraph essay.	challenges of effective communication, equipping them with a versatile set of skills that extend beyond the confines of a standardized writing format.

Essay 1 from Table 6 models the typical response structure to a synthesis prompt.[22] The first body paragraph presents the claim from the first essay, the second presents the counterargument, and the third body paragraph tries to blend them together. The outcome of the third paragraph then shapes the concluding and introductory paragraphs. With that said, ChatGPT may also generate a synthesis structure where each body paragraph merges points from each essay (i.e., essay 2 from Table 6). As with the other essays, ChatGPT does a proficient job with form; students should be given the opportunity to improve its ideas, however.

While ChatGPT produces good form and questionable ideas for students to base their revisions on, it is unclear if students will recognize good structure and poor ideas and if they are open to learning this way. Will students feel like they have learned from ChatGPT outputs and see how they can improve upon ChatGPT's first drafts? If students notice the lack of concrete ideas presented, then ChatGPT can function as a tutor without much oversight, but if students fail to notice this, instructors need to demonstrate that ChatGPT is often providing meaningless and or untrustworthy responses.

6 Student Survey Findings

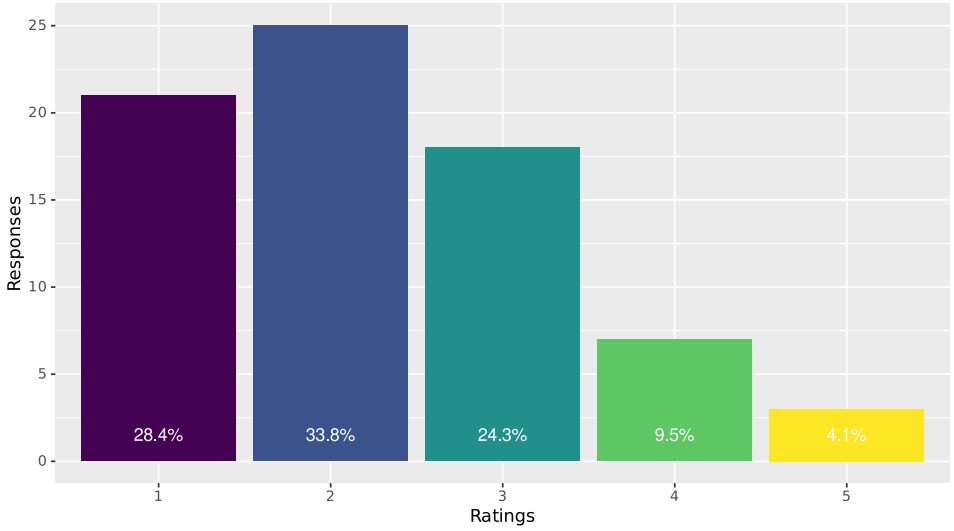
Introductory composition students enrolled at the Florida Institute of Technology (Florida Tech) were given several assignments in which they used ChatGPT to revise and generate texts in the fall of 2023. After completion of these assignments, 74 freshman composition writers at Florida Tech responded to an online survey rating ChatGPT's pedagogical value and

22. This appears to be the most common synthesis output structure.

essay writing performance. The individual survey questions followed a Likert Scale: 1 indicates *Very Good* and *5 Very Poor*.

6.1 Figure 1

Is ChatGPT useful for brainstorming?



Results from Figure 1 show student opinion of ChatGPT's ability to generate new thesis statements on a topic and or classify topics in a new way. More than 50% responded positively.

6.2 Figure 2

How well can ChatGPT model/teach essay structure?

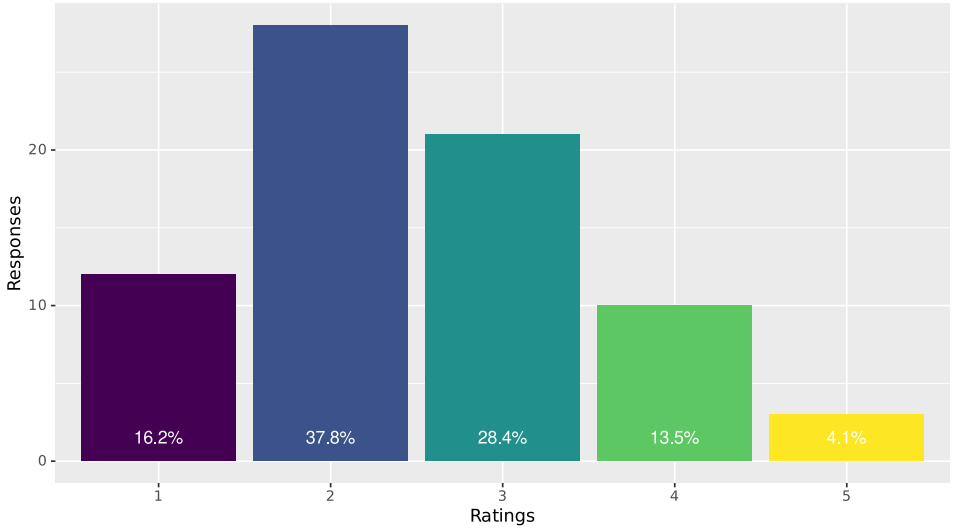


Figure 2 shows student opinion of ChatGPT's ability to model essay structure. While the percentage of 1-responses decreased, still more than 50% responded positively: choices 1 and 2. Figures 1 and 2 show most students felt ChatGPT was useful for brainstorming and modeling basic structure.

6.3 Figure 3

How well can ChatGPT generate synthesis essay structure?

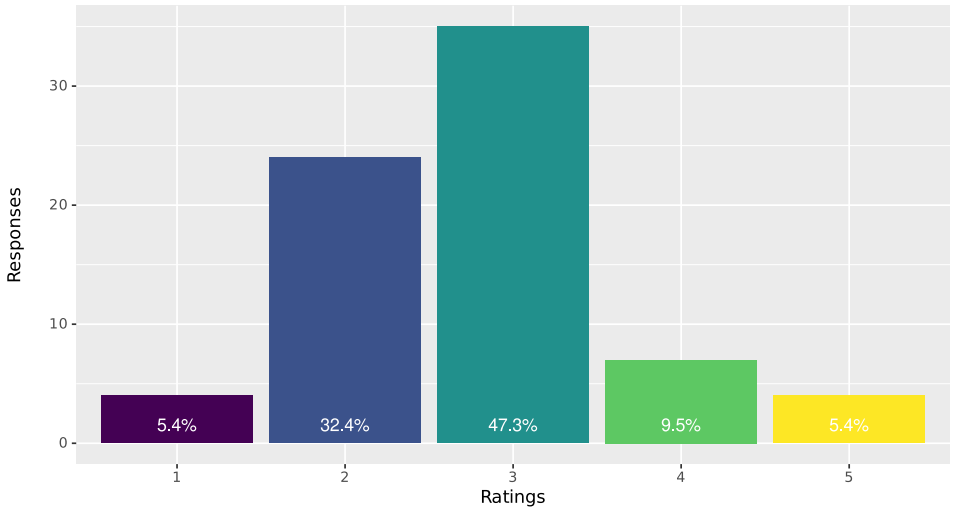


Figure 3 shows student opinion of ChatGPT's ability to model synthesis essay structure. While the percentage of positive responses fell below 50% (very few ranked ChatGPT as highly proficient at generating synthesis structure), only 15% approximately ranked its ability negatively.

6.4 Figure 4

Rate ChatGPT's ability to generate new ideas from synthesis writing.

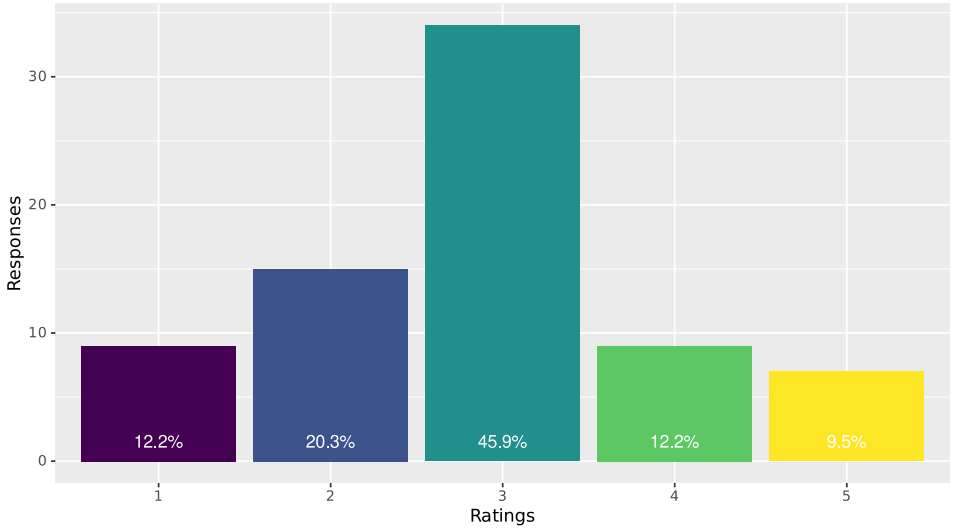


Figure 4 shows student opinion of ChatGPT's ability to generate new ideas from the synthesis process. As in Figure 3, most rated it neutrally (choice 3). While the percentage of choice 1 responses grew, the overall percentage of positive responses decreased.

7 Discussion and Conclusion

This paper explored how to incorporate ChatGPT into writing courses with the goal of using writing to develop higher-order thinking skills. As discussed, these important skills should prepare students for success in higher education (Broad 2003; Lundstrom et al. 2015) and in society (Kravchenko 2009; Kasneci et al. 2023). Hegelian synthesis writing was chosen for its potential to promote higher-order thinking skills. Instructors can use ChatGPT to help students practice the form of argumentative and synthesis writing. Moreover, instructors can use the structure of synthesis writing to prevent students from using ChatGPT to bypass the kind of critical thinking development - a concern of Ye et al. (2024) - that has been (Broad 2003) and should continue to be goal of writing education (Bean,

Melzer 2021). LLMs have only made critical thinking skills more valuable for students (Kasneci et al. 2023).

This study shows students can use ChatGPT to repeatedly generate synthesis essays if they choose an appropriate topic and follow a series of three prompts. Once a viable topic is chosen and inserted into three sequential prompt templates, ChatGPT reliably produces five paragraph argumentative, counter-argumentative, and synthesis essays. Argumentative and counter-argumentative essays follow the traditional structure of introduction, thesis, itinerary, supporting paragraphs, and conclusion. ChatGPT synthesis essays, on the other hand, generally follow the structural pattern of introduction, claim 1, counterclaim, synthesis, and conclusion paragraph. Sometimes, however, ChatGPT produces a structure in which the introduction is followed by three body paragraphs, each discussing a different point of synthesis. Both structures appear helpful for students attempting to map out the form of the Hegelian synthesis process, but neither output can finish the assignment for students when students are challenged to seriously evaluate ChatGPT's ideas.

The results of this study suggest that students cannot use ChatGPT to complete the synthesis process without engaging in critical thinking. Students still need to critically evaluate texts – as discussed in Oostdam (2005) – and draft revisions from feedback, and if students are challenged to create specific, synthesis-derived claims for their synthesis essays, this will push them to go beyond ChatGPT's limitations. While ChatGPT can model and teach the form of writing, it fails at functional competency (Mahowald et al. 2024) and as a truth seeker (Bergstrom, Ogbunu 2023). This study found that its assessment of essays is untrustworthy.[23] This deficit leaves room for student gain. Since LLMs are theoretically limited to the data they are trained on, humans are necessary to evaluate the truth values of premises and to develop viable, innovative solutions. The findings of this paper, thus, suggest that students will engage in critical thinking

23. This finding disagrees with Kasneci et al. (2023) who propose LLMs can be used as an assessment tool.

through the synthesis writing process if instructors make ChatGPT deficits explicit – as Kasneci et al. (2023) suggest doing – and require students to evaluate ChatGPT for bias and vague, meaningless conclusions.

The use of Hegelian synthesis creates opportunities for students to analyze outputs for bias. The point and counterpoint process exposes favoritism when one side is given preference. This was seen with the responses to prompts given in Table 5, which showed GPT's tendency to favor itself. It would not argue against AI without inserting lines acknowledging the benefits of AI. Students can be directed to find similar cases of bias in ChatGTP outputs. With that said, identifying viable topics and ways to help students scrutinize texts for bias is a matter for future research as trial and error is currently the only way to be certain if a topic will work.

The findings of this paper support initiatives to use LLMs in writing education where AI is seen as a tool for brainstorming and creating rough drafts (Martin 2023), and partially support initiatives that see AI as a tool for creating new ideas (Pirehpour 2023). To use ChatGPT to truthfully generate new ideas, methods must be developed to make sure students critically challenge ChatGPT's semantically vacuous responses. This study, thus, serves as work in progress for writing instructors who are looking to integrate AI into their classrooms in accordance with university-wide AI initiatives. ChatGPT can quickly provide templates and ideas for students who have not developed the resilience to struggle for answers. Most writers have difficulty imagining the possibility of a counterpoint to commonly held beliefs and ways of doing things. For example, imagining a counterpoint for the thesis of essay 1 in Table 1 is a difficult task. If asked to develop a counterargument (i.e., provide a reason to abandon/revise the traditional 5-paragraph structure), many students will not know where to begin. They may give up. In such cases, ChatGPT can be prompted to initiate new ways of thinking about problems[24] and staged assignments

24. Essay 2 from Table 1 is a critique of the restrictive nature of academic writing structure. It proposes essay structure should be more flexible. ChatGPT can promptly give students an idea like this, but students will need to use reflective reasoning to evaluate if and how such ideas actually work.

(like synthesis writing) can push students to build from these responses. ChatGPT's help should encourage students to continue while the stages of the assignment slow them down to be more reflective.

The student surveys show that students perceive ChatGPT positively as a brainstorming resource and as a means for learning basic argumentative essay structure. The results indicate that students are open to learning from ChatGPT and recognize the essay structure modeled by it. However, the results show a decline in confidence with synthesis structure. Perhaps more instructor guidance is needed at this stage because students are comparably unfamiliar with this kind of writing. For instance, the results show that fewer students noticed the vacuous nature of ChatGPT's responses. Figure 4 shows only a slight decline in confidence between judgments of ChatGPT's ability to generate synthesis structure (1–2: 37.9%) and to generate meaning (1–2: 32.5%), even though ChatGPT's ability to generate structure is much better than its ability to create new ideas. Students did not appear to notice this difference. Students may need clearer instruction on why ChatGPT's synthesized conclusions are semantically vacuous. Instructors could require students to explain how ChatGPT's synthesized conclusions might work in a world with practical constraints. More work is necessary to improve student outcomes in this area.

In conclusion, AI can help students practice the structure of synthesis writing, and the structure of synthesis writing can help students engage in higher-order thinking. Students still need to develop and insert better ideas into the template generated by ChatGPT. Assignments should challenge them to go beyond the limitations of LLMs. Moreover, the revisions and stages of synthesis writing promote slow thinking, which has been a beneficial part of traditional academic writing. This process affords the opportunity to read critically for bias and to discover and develop new ideas. If we modify writing assignments to develop higher-order thinking skills in the age of LLMs, we can perhaps use LLMs to challenge students to think more deeply about important problems. Since this technology is changing rapidly, further research and refined designs are also necessary

to make this outcome more robust and to preserve writing's place as a vehicle for developing higher-order thinking skills.

Publication Details and Disclosures

Acknowledgments

I want to thank the kind reviewers of this article. Your insights and constructive feedback positively transformed this article. I also want to thank the participants of the *Writing in the Age of Artificial Intelligence* conference in 2023 at the University of Florida. This project was enhanced by your attention and feedback. Thank you!

Funding

The project did not use any grant or specifically allocated research funds. Conference travel funds from the Florida Institute of Technology's School of Arts and Communication were used to help present and subsequently develop the ideas of this article.

Generative AI Use

Generative AI was not used beyond cited text and tables to create the content and or structure of this article.

Biography

Robert Joel Deacon received a PhD in linguistics from the University of Florida in 2014. He has taught composition and research writing at Nagoya University, Structured English Immersion at Arizona State University, and is currently teaching composition and rhetoric at the Florida Institute of Technology. He is interested in the structure of language, writing acquisition, and the relationship between language and cognition.

Copyright

"Developing Thinking through LLM-Assisted Writing: Hegelian Synthesis and Critical Thinking" © 2024 by Robert Deacon is licensed under CC BY-NC-ND 4.0 to the *Journal of Writing and Artificial Intelligence*.

References

ARAPOFF, Nancy, 1967. Writing: A Thinking Process. *TESOL Quarterly*. Online. 1967. Vol. 1, no. 2, p. 33–34. DOI 10.2307/3585751.

BEAN, John C and MELZER, Dan, 2021. *Engaging Ideas: The Professor's Guide to Integrating Writing, Critical Thinking, and Active Learning in the Classroom*. John Wiley & Sons.

BERGSTROM, Carl T and OGBUNU, Brandon, 2023. ChatGPT Isn't 'Hallucinating.' It's Bullshitting. *Undark Magazine*. Online. 2023. Available from: https://undark.org/2023/04/06/chatgpt-isnt-hallucinating-its-bullshitting/

BERLIN, James A, 1996. *Rhetorics, Poetics, and Cultures: Refiguring College English Studies*. Anderson, SC: Parlor Press LLC.

BETTHÄUSER, Bastian A., BACH-MORTENSEN, Anders M. and ENGZELL, Per, 2023. A Systematic Review and Meta-analysis of the Evidence on Learning during the COVID-19 Pandemic. *Nature Human Behaviour*. Online. 2023. Vol. 7, no. 3, p. 375–385. DOI 10.1038/s41562-022-01506-4.

BLOOM, Benjamin Samuel, 1956. *Taxonomy of Educational Objectives: The Classification of Educational Goals*. New York: David McKay Company. ISBN 9780582323865.

BROAD, Bob, 2003. *What We Really Value: Beyond Rubrics in Teaching and Assessing Writing*. Online. Logan: Utah State University Press. ISBN 9780874214802. Available from: https://digitalcommons.usu.edu/usupress_pubs/140/

BURKHALTER, Nancy, 2015. A Dialectical Approach to Critical Thinking Through Writing. *Inquiry: Critical Thinking Across the Disciplines*. Online. 2015. Vol. 30, no. 1, p. 17–23. DOI 10.5840/inquiryct20153014.

CHOMSKY, Noam, 1957. *Syntactic Structures*. Online. The Hague: Mouton & Co. ISBN 9783112316009. Available from: https://www.degruyter.com/document/doi/10.1515/9783112316009/html

CONDON, William and KELLY-RILEY, Diane, 2004. Assessing and Teaching What We Value: The Relationship between College-level Writing and Critical Thinking Abilities. *Assessing Writing*. Online. 2004. Vol. 9, no. 1, p. 56–75. DOI 10.1016/j.asw.2004.01.003.

ELBOW, Peter, 1983. Teaching Thinking by Teaching Writing. *Change: The Magazine of Higher Learning*. Online. 1983. Vol. 15, no. 6, p. 37–40. DOI 10.1080/00091383.1983.10570005.

ELBOW, Peter, 1994. Teaching Two Kinds of Thinking by Teaching Writing. *Re-thinking reason: New perspectives in critical thinking*. 1994. P. 25–32.

EMIG, Janet, 1977. Writing as a Mode of Learning. *College Composition and Communication*. Online. 1977. Vol. 28, no. 2, p. 122–123. DOI 10.2307/356095.

FRANKFURT, Harry G., 2005. *On Bullshit*. Online. Princeton, NJ: Princeton University Press. ISBN 9780691122946. Available from: https://press.princeton.edu/books/hardcover/9780691122946/on-bullshit

HALLORAN, Clare, JACK, Rebecca, OKUN, James and OSTER, Emily, 2021. *Pandemic Schooling Mode and Student Test Scores: Evidence from US States*. Online. Available from: http://www.nber.org/papers/w29497.pdf

HARRIS, Roy, 2009. Speech and Writing. In: OLSON, David R. and TORRANCE, Nancy (eds..), *The Cambridge Handbook of Literacy*. Online. Cambridge: Cambridge University Press. p. 46–58. Cambridge Handbooks in Psychology. ISBN 9780521862202. Available from: https://www.cambridge.org/core/books/cambridge-handbook-of-literacy/speech-and-writing/354B3E7A94BCAC7DF0B76E266A2ACC41

KAHNEMAN, Daniel, 2011. *Thinking, Fast and Slow*. Online. 1st ed. New York: Farrar. ISBN 9781429969352. Available from: https://us.macmillan.com/books/9780374533557/thinkingfastandslow

KASNECI, Enkelejda, SESSLER, Kathrin, KÜCHEMANN, Stefan, BANNERT, Maria, DEMENTIEVA, Daryna, FISCHER, Frank, GASSER, Urs, GROH, Georg, GÜNNEMANN, Stephan, HÜLLERMEIER, Eyke, KRUSCHE, Stephan, KUTYNIOK, Gitta, MICHAELI, Tilman, NERDEL, Claudia, PFEFFER, Jürgen, POQUET, Oleksandra, SAILER, Michael, SCHMIDT, Albrecht, SEIDEL, Tina, STADLER, Matthias, WELLER, Jochen, KUHN, Jochen and KASNECI, Gjergji, 2023. ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education. *Learning and Individual Differences*. Online. 2023. Vol. 103, no. 102274, p. 102274–102275. DOI 10.1016/j.lindif.2023.102274.

KRATHWOHL, David R., 2002. A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*. Online. 2002. Vol. 41, no. 4, p. 212–218. DOI 10.1207/s15430421tip4104_2.

KRAVCHENKO, Alexander V., 2009. The Experiential Basis of Speech and Writing as Different Cognitive Domains. *Pragmatics & Cognition*. Online. 2009. Vol. 17, no. 3, p. 527–548. DOI 10.1075/pc.17.3.03kra.

LAI, Emily R, 2011. Critical Thinking: A Literature Review. *Pearson's Research Reports*. 2011. Vol. 6, no. 1, p. 40–41.

LAI, Wai Ling, 2017. The Secret of Building a Logical Argument. *NU Ideas*. 2017. No. 6, p. 1–10.

LIBERMAN, Alvin M, 2017. The Relation of Speech to Reading and Writing. In: DE GELDER, Beatrice and MORAIS, Jose (eds..), *Speech and Reading: a Comparative Approach*. London: Taylor, Francis. p. 17–31. Psychology Library Editions: Psychology of Reading. ISBN 9781315111810.

LUNDSTROM, Kacy, DIEKEMA, Anne, LEARY, Heather, HADERLIE, Sheri and UNIVERSITY, Utath State, 2015. Teaching and Learning Information Synthesis: An Intervention and Rubric Based Assessment. *Comminfolit*. Online. 2015. Vol. 9, no. 1, p. 60–61. DOI 10.15760/comminfolit.2015.9.1.176.

MAHOWALD, Kyle, IVANOVA, Anna A., BLANK, Idan A., KANWISHER, Nancy, TENENBAUM, Joshua B. and FEDORENKO, Evelina, 2024. Dissociating Language and

Thought in Large Language Models. *Trends in Cognitive Sciences*. Online. 2024. Vol. 28, no. 6, p. 517–540. DOI 10.1016/j.tics.2024.01.011.

MALEKI, Negar, PADMANABHAN, Balaji and DUTTA, Kaushik, 2024. AI Hallucinations: A Misnomer Worth Clarifying. *arXiv.org*. Online. 2024. Available from: https://arxiv.org/abs/2401.06796v1

MARTIN, Juzanne, 2023. Exploring Artificial Intelligence's Impact and Role in College. *Pegasus*. Online. 2023. Available from: https://www.ucf.edu/pegasus/what-can-ai-teach-us/

MENARY, Richard, 2007. Writing as Thinking. *Language Sciences*. Online. 2007. Vol. 29, no. 5, p. 621–632. DOI 10.1016/j.langsci.2007.01.005.

OATLEY, Keith and DJIKIC, Maja, 2008. Writing as Thinking. *Review of General Psychology*. Online. 2008. Vol. 12, no. 1, p. 9–27. DOI 10.1037/1089-2680.12.1.9.

OOSTDAM, Ron, 2005. Assessment of Argumentative Writing. In: RIJLAARSDAM, Gert, VAN DEN BERGH, Huub and COUZIJN, Michel (eds..), *Effective Learning and Teaching of Writing*. Online. Dordrecht: Springer Netherlands. p. 427–442.

OPENAI, 2023. ChatGPT. *3.5*. Online. 2023. Available from: https://chatgpt.com/

PETITTO, Laura Ann and MARENTETTE, Paula F., 1991. Babbling in the Manual Mode: Evidence for the Ontogeny of Language. *Science*. Online. 1991. Vol. 251, no. 5000, p. 1493–1496. DOI 10.1126/science.2006424.

PIREHPOUR, Kevin, 2023. ASU Students Explore Ways to Embrace AI Tools in Writing Course. *ASU News*. Online. 2023. Available from: https://news.asu.edu/20230523-asu-students-explore-ways-embrace-ai-tools-writing-course

PREISS, David D., CASTILLO, Juan Carlos, GRIGORENKO, Elena L. and MANZI, Jorge, 2013. Argumentative Writing and Academic Achievement: A Longitudinal Study. *Learning and Individual Differences*. Online. 2013. Vol. 28, no. 2013, p. 204–211. DOI 10.1016/j.lindif.2012.12.013.

SOUTHWORTH, Jane, MIGLIACCIO, Kati, GLOVER, Joe, GLOVER, Ja'Net, REED, David, MCCARTY, Christopher, BRENDEMUHL, Joel and THOMAS, Aaron, 2023. Developing a Model for AI across the Curriculum: Transforming the Higher Education Landscape via Innovation in AI Literacy. *Computers and Education: Artificial Intelligence*. Online. 2023. Vol. 4, no. 100127. DOI 10.1016/j.caeai.2023.100127.

YE, Andre, MOORE, Jared, NOVICK, Rose and ZHANG, Amy X., 2024. Language Models as Critical Thinking Tools: A Case Study of Philosophers. *arXiv.org*. Online. 2024. Available from: https://arxiv.org/abs/2404.04516v1