

# Critiquing ChatGPT Compositions: Collaborative Annotation as an Approach to Enhancing Students' Metalinguistic Awareness of AI-Generated Writing

Ruth Li

## ***Abstract***

*In this article, I introduce a collaborative annotation activity that supports students in critically examining AI-generated writing in relation to criteria including specificity and complexity. I engage students in collaboratively annotating the AI-generated essays, guiding students to identify instances in which the essays could be more specific—clearly defined or identified—and complex—considering various perspectives or aspects of an issue. I then immerse students in reflecting on how the essays could be revised to be more specific and complex. I posit that this exercise sharpens students' metalinguistic awareness of AI writing. By attending closely to AI-generated essays, students critically interrogate AI writing and invite the construction of more nuanced, complicated ideas. Moreover, this exercise enhances students' agency, recasting the learning process as a collaborative, dynamic endeavor that fosters dialogic interactions between humans and AI writers. Ultimately, I open pathways toward more accessible approaches to supporting students' critical engagements with AI.*

**Keywords:** collaboration, annotation, metalinguistic awareness, critical thinking, feedback

## **Introduction, Background, and Context**

The advent of generative AI in education invites writing studies scholars and instructors to rethink approaches to supporting students' development as writers. AI-generated writing has been critiqued as standardized and "flattened" (MLA-CCCC Joint Task Force on Writing and AI, 2023), yet students may be less critically attuned to the potential limitations of AI texts. The increasing prevalence of AI in the classroom raises a need for instructors to develop strategies for enhancing students' critical digital literacy in relation to large language models including ChatGPT (Anderson, 2023), including approaches that inspire students to interrogate the standardized forms of writing that may be perpetuated by AI language models. Such standardized forms emphasize grammatical correctness and disciplinary and genre-based conventions. Scholarship in writing studies (Milroy & Milroy, 2012) has critiqued the historical focus of writing instruction on prescriptivism—"approaches that are based on stipulating prescriptive rules and correcting students' use of grammar" (Thieme, 2022, p. 281). Tremain et al. (2023) posit that current approaches to teaching discipline-specific writing continue to "uphold entrenched disciplinary genres, particular criteria for 'good' writing, and certain beliefs about style and grammar" (p. 121). In responding to the pedagogical challenges of supporting students to engage critically with AI-produced writing, I invite spaces for students to critique and complicate conventional forms of writing that emphasize

structure and grammar, and to experiment with innovative possibilities of expression and meaning-making. In this article, I offer an approach that encourages students' careful attention to the ideas and language of AI-generated texts.

In this article, I introduce a collaborative annotation activity that supports undergraduate students in critically examining AI-generated writing in relation to criteria including specificity and complexity. I designed and taught this assignment in spring 2024 as part of a new upper-level undergraduate elective course on communication in the age of artificial intelligence and digital media. This course contributes to a minor in strategic communication offered by the English department at the university where I teach. The course description, which I developed alongside the department chair, appears below:

New technologies are constantly making our communication tasks both easier and more complicated. Students will learn about a variety of AI-driven communication tools, including ChatGPT, social media algorithms, speech analysis, chatbots, and many other tools. Interactive class discussions as well as firsthand exploration of these tools will help students understand the incredible power as well as the challenges and limitations of using artificial intelligence and digital media for communication. Students will have opportunities to analyze AI-generated texts, to design an AI innovation, and to compose using multiple media forms, including visual, audio, and linguistic components.

The annotation activity I describe below supports the course objectives: to examine the affordances and limitations of artificial intelligence tools, to interpret the substance and style of human-written and AI-produced texts, and to understand how we can best use AI and digital media tools to communicate. I teach at a university in Saudi Arabia, where the majority of the students are multilingual and come from diverse cultural backgrounds. While the university is an English medium of instruction institution, many students learn English as a second language in addition to Arabic. The majority of students in the class are junior and senior business and engineering students who bring their disciplinary interests to the class discussions. For instance, a student interested in entertainment marketing chose to annotate a ChatGPT-generated paragraph about musicals, while students majoring in engineering chose to analyze ChatGPT-generated case studies. The results of the activity can thus differ depending on the students' areas of focus. In incorporating this approach within a multilingual context, my objectives are to stimulate students' critical thinking about AI-generated texts. While I offer this course within a writing and communication context, this activity could be adopted by other disciplines that may use writing, including in exercises that encourage students to interrogate ideas or probe assumptions in AI-generated texts across disciplinary contexts.

### **Educational Theories & Frameworks**

This collaborative annotation approach draws from theories of collaboration as “conversation” (Bruffee, 1984). I conceptualize collaborative annotation as grounded in the notion that “writing is a social artifact, like the thought that produces it;” students' annotations could be conceived as a “conversational exchange among peers” (Bruffee, p. 642). This strategy builds on pedagogical approaches to collaborative social annotation, which Kalir (2020) defines as “a genre of learning technology that can productively enable group collaboration and shared meaning-making” (p. 245). Studies have illustrated the ways social annotation can promote domain-specific knowledge

and inquiry (Zhu et al., 2020), support the construction of knowledge (Morales et al., 2022), and facilitate collaboration among peers (Chan & Pow, 2020). In extending social annotation approaches to the era of generative AI, I examine how collaborative digital annotation, which can enable synchronous interaction in a shared space (Barton & Klint, 2011), can support students' critical engagements with AI-generated texts. I conceptualize *collaboration*, then, in at least two senses: as students' interactions with one another and as students' interchanges with AI-generated texts.

In conceptualizing interactions between students and AI as collaborative, I acknowledge a tendency to anthropomorphize AI. Anthropomorphism is defined as “the process of attributing humanlike motivations, emotions or characteristics to real or imagined non-human entities” (Alabed et al., 2022, p. 5). As Alabed et al. explain, chatbots such as ChatGPT have the capacity to produce human-like conversation, while AI customer service agents or mental health counselors might display personalities, appearances, or emotions including warmth and empathy. While I acknowledge the ways the physical or affective attributes of AI tools may lead users to anthropomorphize the tools, I focus on the textual dimensions of composing with AI. I conceptualize human-AI interactions as collaborative in the sense that human writers can engage in an unfolding dialogue with AI tools: writers can critically examine and evaluate AI-generated essays, identifying areas that could be more specific and complex. Writers can then prompt ChatGPT to revise an essay and to compose more nuanced, complicated writing. I envision composing with AI as a processual endeavor in which human writers can critique, complicate, or expose gaps or elisions in AI-generated texts while stimulating the tool to compose more nuanced responses (Li, 2024). The exercise could thus be conceptualized as “collaborative” in the sense that the iterative, recursive cycles of feedback and revision could invite a dialectical interplay between human and artificial interlocutors. By investigating students' collaborative annotations, I illuminate the ways students co-construct meanings in their dialogic interactions with peers and AI. In this sense, I re-envision artificially (re)generated writing as a dynamic, interactive process inflected by human writers' critical commentaries and AI responses.

In developing this approach, I also draw from the scholarship on writing feedback, including peer feedback on writing. In their annotations, students comment on the ways AI-generated essays could be revised in relation to criteria including specificity and complexity; I thus conceptualize students' annotations as a form of feedback on AI-generated writing. Research has demonstrated the ways peer feedback facilitates learning (Li et al., 2010) and emphasized the importance of scaffolding peer review in ways that encourage constructive feedback (Kim, 2015). In their study of collaboration in writing and writing feedback, Wigglesworth & Storch (2012) posit that working together on writing tasks “can enhance the language learning opportunities for learners through their ability to scaffold each other's contributions and knowledge” (p. 364). Elucidating Wigglesworth and Storch's findings, my analysis below illustrates the ways students' annotations build on and complicate one another as students formulate more critical commentaries over the course of the annotation process.

In addition, this collaborative annotation activity aims to foster students' metacognition, or “thinking about thinking” (Flavell, 1979). By encouraging students to attend closely to specific features of AI-generated essays, including ideas, structure, and language, I strive to enhance students' metalinguistic awareness of writing, stimulating students' close, careful attention to the intricacies of language as it creates meaning in discourse. Metalinguistic awareness is defined as “the ability to reflect on and manipulate the structural features of language” (Nagy et al., 1995, abstract). Engaging students in collaboratively critiquing the linguistic features of AI-generated

writing can render visible the learning process and stimulate an “explicit and conscious knowledge of language” (Roehr-Brackin, 2018, abstract). An “explicit attention to language itself” (Schleppegrell, 2013, p. 156) can be especially useful for multilingual learners and L2 students, who may need additional support with academic language development. Moreover, with the rise of AI in classroom contexts, there emerges a need to support students to critically consider the affordances and limitations of AI-generated writing. In synthesizing the scholarship in collaboration, social annotation, feedback, and metacognition, I contribute an innovative approach to collaborative annotation in the era of generative AI. My aims are to support students with critically examining AI-generated writing.

## Approach

I describe the collaborative annotation approach in detail below. I incorporated this exercise toward the beginning of the semester as part of a unit that immerses students in interrogating the extent to which AI-generated texts could be conceptualized as “creative,” though the activity could be adapted for a range of curricular contexts, including assignments that encourage students to investigate the ways AI-generated content is constructed across disciplinary domains. In the guidelines, I ask students to work in small groups (2-3 members each) to annotate (add comments on) a ChatGPT-generated essay. To generate example AI-generated texts for analysis, I prompt ChatGPT (version 3.5) to compose essays on specific topics, for instance, on the affordances and limitations of AI or the benefits and disadvantages of social media. I engage students in collaboratively annotating the ChatGPT-generated essay on a shared Google Doc using the “comment” function. While I used Google Docs for this activity, other social annotation platforms, such as Hypothesis or Perusall, could also be used as tools to encourage collaborative commentary on a shared text.

I guide students to identify instances in which the essays could be more *specific*—clearly defined or identified—and *complex*—considering various perspectives or aspects of an issue. In particular, I encourage students to consider questions such as:

- How specific, complex, or nuanced are the ideas in the essay?<sup>1</sup> Where are places that could be more specific, complex, or nuanced? Which ideas could be developed further? Where could examples be added? Which alternative perspectives could be considered? To what extent does the essay present original or surprising ideas?

The full guidelines for the exercises appear in the appendix. To scaffold the annotation process and explain what is meant by criteria including specificity and complexity, I present to students an example annotation in which I comment on a sample student-written thesis statement. In the example comment, I note places where the statement could be revised to be more specific and complex (e.g., “Specify the benefits and harms?”). Alongside the original version of the thesis statement, I include a revised version with elements of specificity highlighted in green and elements of complexity highlighted in blue. The example annotation can be accessed in the resource below:

---

1. Elsewhere (forthcoming, in progress) I have written in more detail regarding the conception of *nuance* and approaches to supporting students to write in nuanced ways that signal a recognition of the subtle shades or variations of meaning and expression. In this class’s unit on AI and creativity, I immersed students in discussing what the word “nuance” meant to them and in critiquing the extent to which AI-generated texts, including poetry and screenplays, could be conceived as nuanced.

*Resource 1: Modeling an example annotation*



**Students' Collaborative Annotations**

To illustrate this collaborative annotation approach in action, I analyze example student annotations and written reflections. The following resource shows students' collaborative annotations on a ChatGPT essay that examines the pros and cons of social media<sup>2</sup>:

*Resource 2: Students' collaborative annotations on an ChatGPT-generated essay*



Highlighting the phrase “pros and cons of social media” in the generated essay, one student notes, “positives and negatives,” while another student comments, “A strong thesis statement.” The student’s positive evaluation of the thesis statement contradicts the guidance provided in the example annotation, which critiques the ways the example thesis statement could be revised to be more specific and complex. In counterpoint with the second student’s positive evaluation, a third student offers a more critical commentary: “This thesis statement could be even stronger if those pros and cons were alluded to or mentioned specifically rather than generally.” In drawing from the guid-

---

2. This was the exact prompt I used to generate the sample essay: “Write an essay on the pros and cons of social media.” I acknowledge that the generic nature of the prompt may have led to more generalized outputs; in future iterations of this activity, I might experiment (and encourage students to experiment) with more specific, tailored prompts.

ance provided in the instructions, the student recognizes the generalized nature of the thesis statement and identifies a need for greater specificity in the claims presented.

The students' annotations in response to the second paragraph of the AI-generated writing identify the paragraph's elements, including statements of argument and support. Highlighting the sentence "...social media facilitates connectivity and communication on a global scale...", a student comments, "states an argument." Meanwhile, in response to the sentence "Social media platforms also serve as powerful tools for information dissemination, enabling rapid sharing of news, ideas, and resources," a student comments, "backs up argument." While the students' comments on the second paragraph offer relatively neutral observations on the essay's elements, the students' annotations on the third and fourth paragraphs critically evaluate the essay in relation to criteria including specificity and elaboration: one student writes, "The AI is very direct and straightforward, but it lacks further elaboration and details," while another writes, "The AI is very specific about privacy concerns, but it does not provide examples related to the potential privacy concerns of social media to back its argument." By attending closely to its composition, the students critique the affordances and limitations of the essay in a balanced manner, acknowledging the directness of the writing while noting the lack of examples and support. Interestingly, the students' comments on the lack of "elaboration and details" in the ChatGPT-generated essay echo Brown & Aull's (2017) findings that lower-scoring Advanced Placement essays are characterized by "emphatic generality," while higher-scoring essays illustrate "elaborated specificity," or "focusing on and explicating specific, often abstract, concepts" (p. 394). In this sense, the students' comments illuminate salient linguistic characteristics of writing, which range on a spectrum from generality to specificity, from emphasis to elaboration.

For the second part of this activity, I invited students to prompt ChatGPT (or another tool of their choice) to write an essay (or a report, a case study, an abstract, etc.) on a topic of their choice. I asked students to copy and paste the generated text into the shared Google Doc and to annotate the AI-generated text based on the guiding questions above, including by noting places where the text could be more specific or complex. As illustrated in part II of resource 2 linked above, one group of students prompted ChatGPT to write a paragraph about musicals. Commenting on the first sentence of the generated text, a student writes, "Doesn't seem to have a thesis statement," a critique that carries normative assumptions about what the elements of an essay. Such assumptions may stem from the model annotation, which comments on a sample thesis statement. In addition, the students highlighted a sentence that mentions examples of musicals, commenting: "This can be seen as case studies or examples, even written comparatively with classics and contemporary examples." By referring to the examples as "case studies," the students signal their recognition of the specific writing genres that the generated piece emulates. The students also highlighted the word "indelible," noting that it is "a word used commonly in AI text, which takes away from its originality. Interestingly, the students' observation corroborates the "common" usage of the word with a lack of "originality," implying that overused expressions may be seen as less original. The students' observations echo Markey et al.'s (2024) findings that "generated texts are more informationally dense than authored texts and often read as dialogically closed, 'empty,' and 'fluffy'" (n.p.)—lacking in substance and distinctiveness.

I also encouraged students to prompt the tool to revise the AI-generated text based on their comments or critiques and to annotate the revised text, though the students did not initiate further revisions. I then immersed students in a class discussion on what they noticed about the generated texts, including ways the texts could be revised to be more specific, complex, and nuanced.

## Students' Individual Annotations & Reflections

As part of the course's midterm, I instructed students to individually annotate a ChatGPT-generated essay<sup>3</sup> on the affordances and limitations of AI and to write a short paragraph-length response that reflects on what they noticed about the essay and identifies ways the essay could be revised to be more specific and complex. A student's annotations can be accessed in the resource below:

*Resource 3: A student's individual annotations on an ChatGPT-generated essay*



The student offers thoughtful, astute critical commentary on the essay's construction, critically examining issues including coherence, exemplification, and the development of ideas. For instance, the student comments on the cohesiveness and flow of the piece ("Grammatically, I'm not sure if this works. Reading this doesn't flow as easily as an authentic text;" "It seems to be jumping between ideas without a smooth transition between them, or much detail provided about why these affordances are important"). In the first comment, the student draws an interesting distinction between an "authentic" text and an AI-generated text, observations that are further developed in the student's reflection, which appears below. In following the guidelines in the prompt, the annotations shift between statements of critical evaluation ("This is a bit general") and constructive commentary ("A stronger thesis statement would go into more detail regarding the potentials and constraints to allow readers to have a better idea of what they're about to read"). In this comment, the student draws a contrast between generality and specificity as indicators of writing quality. The student also comments on the scope of the claims ("This is a big claim to make without any examples or evidence") and the underdevelopment of ideas ("a half-thought out idea rather than a fully realized one"). Moreover, the student offers normative statements that reflect rule-based assumptions about what an essay "should" include ("The concluding statement of a paragraph should summarize the points made within that paragraph"). Shifting between statements of observation, evaluation, and suggestion, the student's comments resemble instructors' feedback on student papers, echoing the way instructors evaluate the effectiveness of the writing based on a set of criteria. The dialogic nature of the student's annotations is especially evident in the instances in which the student poses questions that encourage elaboration and specification ("Such as? Not much detail is provided about what kind of social inequalities it perpetuates. Is it regarding race, gender, social status, etc?"). In this instance, the student critically interrogates the ChatGPT-generated essay in a

---

3. I prompted ChatGPT 3.5 to "write an essay on the affordances and limitations of AI."

dialectical fashion, probing into its gaps and fissures, exposing the elision of detail regarding the construct of social inequalities. In its disentangling of the text's construction, the student's annotations could be conceived as deconstructionist in nature, seeking to subvert or complicate AI-generated content. In a dialectical interplay between human and AI interlocutors, the annotation process stimulates the construction of more nuanced ideas and perspectives.

The student's written reflection<sup>4</sup> elaborates even more thoroughly on the critiques posed in the annotations:

This essay, though enlightening, faces many of the problems commonly experienced with AI-generated texts. Texts like these should be used as a starting point for generating ideas rather than the final product, as there are clear structural issues within the piece that have altered the end meaning of the essay, which is why it should be argued that texts like these shouldn't be valued as much, due to grammatical issues, general statements without examples to prove a point, as well as introducing new concepts in concluding statements, going against many important rules of writing.

In this reflection, the student delineates the "structural issues" identified in the annotation. The student's conceptions of the "rules of writing"—a clearly defined thesis statement, claims supported by evidence, a conclusion that summarizes the main points of the essay rather than introducing new ideas—are perhaps influenced by the model annotation I shared with the class, in which I lifted a thesis statement from a sample essay that presents the elements of a five-paragraph essay. Interestingly, the student comments on the way the "value" of AI-generated texts diminishes when lacking in such qualities of writing, an idea that is developed further later in the reflection:

Additionally, the generalization is without evidence to support the claim. This can be seen throughout the essay, however, an example of this would be, "Perpetuating social inequalities." Since this is a very sensitive issue to many, making a claim such as this one without evidence shows a lack of detail and care for human matters such as this one. Since AI is a combination of data, there is a sense of apathy that is quite off-putting to readers, creating a separation, which is especially seen in this example as this is a very human issue. This also goes against the basic field of essay writing, in which you have to back up your claims with either evidence or further explaining claims to make it a fully realized idea, which this AI text seems to lack.

The student raises several interesting notions, including the way the AI text "creates a separation" between the reader and the text. From the student's reflection, one could conceptualize a relation between intimacy and distance, connection and disconnection in which reading is construed as an intimate encounter with a text. In this conception, the absence of the human dimension, the absence of feeling ("a sense of apathy") fragments the connective tissue that enjoins the reader and the text, creating an effect that is "off-putting to readers." Such a sense of the "separation" between the reader and the text likewise emerges from the students' conclusion that AI-generated writing is "leaps and bounds behind authentic writing," a statement that associates human writing with "au-

---

4. This project was approved by the institution's IRB (20244). I am grateful to my student Abdulrahman Al Aifan for permission to publish this reflection. It might be noted that the student's reflection goes far beyond my stated expectations for a paragraph-length response.



thenticity,” thereby constructing a duality between authenticity of human writing and the inauthenticity of artificially generated writing. In this reflection, the student offers thoughtful, provocative ideas on the nature of AI-generated writing. A space opens for further interrogating notions such as “authenticity”: in a follow-up discussion, instructors might engage students in disentangling conceptions of the “real,” genuine, or original as implicated by artificially generated texts. In particular, the student’s insights bring to mind Benjaminian conceptions of the “aura” of a text in its mechanical (or digital or artificial) reproduction. Such a sense of the loss of “aura” in the AI-generated essay seems resonant with what the student describes as the “sense of apathy” that emerges from encountering the essay. The student’s ideas echo notions that AI-generated content could “potentially dilute the emotional resonance that often comes from human-generated content, leading to a sense of disconnection” (Tsao & Nogues, 2024, p. 10). In this sense, the student reaches toward theorizing the affective (or disaffective) dimensions of the human-AI encounter, tracing the tensions that emerge from the interface of the human and the artificial.

### **Fostering Metacognitive Discussions on Writing**

One important consideration is that as large language models evolve, generative artificial intelligence may be able to produce more nuanced and complex writing as the technology progresses. Such technological advances raise questions over the ways instructors might facilitate metacognitive<sup>5</sup> conversations on writing in an era in which distinctions between human and AI-generated writing may become increasingly blurred. In engaging students who find AI-generated writing to be near-flawless, instructors might ask students to consider characteristics such as the originality of the ideas, the distinctiveness of the style, or the degree to which the text invites unexpected or surprising insights or complicates existing conceptions. Yet qualities such as subtlety or nuance may defy an analysis of the linguistic elements of a text, transcending its structural features. I thus aim to inspire students to delve into those facets which emerge beyond language: to investigate whether there is an ineffable quality of writing that cannot be captured by artificial intelligence, such as an undercurrent of feeling or a layering of meaning. In fostering discussions on the qualities of writing that may be less readily identifiable in a linguistic analysis, instructors can inspire students to explore the felt sense or experience of a text—the impression a text leaves on a reader. Classroom discussions could delve into the experiential dimension of writing—the presence of consciousness, subjectivity, or voice; the extent to which the text conjures vivid descriptions or creates emotional resonances with the reader. Instructors might engage students in examining questions such as: Which qualities of writing constitute voice? In what ways can we register subjectivity in a text? In what ways does the text construct a sense of intimacy with or distance from a reader? In addressing these questions, students might consider elusive aspects such as the contours of feeling infused into a piece or the depth and dimension of the writer’s perspective.

To discern facets of writing that elude recognition or description is to trace a distinction between thought and expression, to open spaces for the inexpressible or inarticulable. In the essay

---

5. I am inspired by the prefix *meta-* meaning “after” or “beyond”: in stimulating dialogue on AI-generated writing, we might seek beyond the text itself into the underlying processes of its construction: the cognitive processes of critical thinking and writing, the mechanisms by which large language models generate increasingly sophisticated text using predictive sequences, the processes of labor and extraction.

“I type, therefore I am,” Chatfield questions whether “words alone could negotiate the world on our behalf”:

In an era when the gulf between words and world has never been easier to overlook, it’s essential to keep alive a sense of ourselves as distinct from the cascade of self-expression; to push back against the torrents of articulacy flowing past and through us.

To distinguish the self from its articulacy is to seek beyond the texts that have circulated across time and become reified into the corpus: to strive toward the ineffable spaces of being, the capacity for wonder. Paradoxically, to disentangle the possibilities and limits of expression is to linger in the edges of the expressible. Yet the image of a “cascade of self-expression” simultaneously inspires a conceptualization of writing as a process of mediation — as the limn enjoining writer, reader, and text. In theorizing the ways we negotiate meaning in artificially generated texts, we can interrogate new forms of mediation at the interface of the human and artificial — an interface brimming with possibilities of (re)generation.

### **Insights & implications**

This collaborative annotation exercise enables students to visualize learning, rendering visible the annotation process as it unfolds across students’ engagements with AI-generated texts. While annotating collaboratively, students built on one another’s critiques, offering more critical evaluations of the text over the course of the annotations. Students’ individual annotations in the midterm reveal a greater depth of engagement and reflection than in the initial in-class activity, indicating that the collaborative annotation activity served as a scaffold for students’ individual annotations. In stimulating students’ metalinguistic awareness of writing in ways that spark critical engagements with AI outputs, the activity encourages students to attune more closely to fine-grained elements of language in AI-generated texts, including word choices and phrasings, and to consider the effects these features create on a reader. Moreover, this activity enhances students’ agency as knowledge constructors, recasting the learning process as a collaborative, dynamic endeavor (Li, 2024; Al-khresheh, 2024). By positioning students as critical interlocutors of AI-generated texts, the exercise invites participants to “re-think how knowledge is produced” and to engage in a “reciprocal dialogue” with AI-produced writing (Tsao & Nogues, 2024, p. 2). Such a dialogue blurs the boundaries between authors and readers while opening avenues for novel forms of human-AI collaborative writing. Taken together, the students’ annotations illuminate a dialectic interplay between human and AI interlocutors: students expose and interrogate the gaps and elisions in the AI-generated essays while opening spaces for developing and refashioning AI writing. Yet students’ annotations reveal interactions not only with AI texts but also with one another: students’ annotations complicate one another, reaching toward more critical examinations of ChatGPT writing. The collaborative annotations illustrate a developmental trajectory as students deepen and expand their insights over the course of the annotations and reflections. In this sense, the exercise indicates the potential for the collaborative construction of critique, illuminating the ways students construct critical commentaries of AI-generated texts through dialogic interactions with peers.

At the same time, the outcomes of the annotation activity reveal a need to further support students’ critical thinking and knowledge construction in relation to AI-generated writing. Stu-

dents' uncritical commentary on the generalized statements in the AI-generated essays raises implications for supporting students to critically interrogate AI-generated writing in relation to criteria including specificity, while students' conceptions of the "rules of writing" open avenues for encouraging students to question rule-based assumptions of writing and to recognize the rhetorically situated nature of knowledge construction across generic and disciplinary domains. A few limitations of the activity include the narrow, decontextualized language used in the model annotation (e.g., the focus on generic elements of essay writing such as "thesis statements") and the generic prompts I used to generate the essays. In future iterations of this activity, I plan to create more specific, tailored prompts that frame the parameters of the output, including the content and structure, in relation to particular generic and disciplinary domains. In addition, the multilingual context of my classroom posed particular challenges, as students were immersed in learning English communication skills while also learning to engage in annotation exercises that were unfamiliar to some students; a few students struggled with understanding the instructions and grasping what it meant to annotate a text. To enhance students' critical thinking while supporting their academic language development, I aim to offer more scaffolding, including modeling and practice with the annotation process. The exercise thus invites broader implications for supporting L2 pedagogy in the context of AI.

## **Practical Guidance & Recommendations**

- Provide models and scaffolds for collaborative annotation.

To support students' critical engagement with AI texts, it is important to scaffold the annotation process. For instance, teachers could provide models of student- or teacher-created annotations and ask students what they notice about the ways the annotations comment on or critique the text. Offering collaborative and individual practice as well as choice of text, topic, or prompt could also enhance students' learning outcomes. Another possible area for development could be to sharpen students' attention to the particular criteria they are evaluating. While my students were able to identify instances where the AI-generated essays could be more specific, instructors could develop ways to scaffold students' recognition of complexity — the multiple components of a text. By comparing human-written and AI-generated writing in a similar genre, for instance, students could learn to recognize salient distinctions between human and artificial texts and to identify ways in which AI-generated essays might invite alternative perspectives or engage with the nuances of an issue.

- Engage students in the process of prompt engineering, or crafting specific prompts.

As students noted in class discussions, the relative specificity or generality of the prompt could influence the quality of the output; as one student noted in their reflection, "Perhaps if the prompt was more specific these issues [with the AI-generated text] would be resolved." Knoth et al. (2024) define prompt engineering as "formulating precise and well-structured instructions" (p. 1). In prompting AI to generate outputs, users could specify the parameters of the prompt, including genre and task: for instance, students could prompt ChatGPT to compose a business case study that examines users' experiences with AI. Instructors could immerse students in iteratively refining prompts and analyzing the relative effectiveness of each successive iteration. By engaging students in crafting more tailored, targeted prompts, instructors could inspire students to conceptualize

knowledge construction as an iterative, recursive, and multilayered process of refining meaning (Li, 2024).

- Experiment with various AI and annotation tools, including Gemini, Hypothesis, Perusall, and Canvas.

One interesting consideration for this activity is Google's move to integrate Gemini into Google Docs. Using Google Gemini in Google Docs, users can prompt the tool to "write new text" or "rewrite existing text": "for example, you can rephrase text, or you can make it more formal, more concise, or more detailed" (Google Docs Editors Help). The integration of Gemini into Google apps could potentially contribute to this assignment: for instance, students could ask Gemini to critique generated essays; students could then add a layer of metacommentary by reflecting on the effectiveness of Gemini's suggestions. As Google describes, options for rewriting a text using Gemini include:

Tone: Select Formal or Casual  
Summarize: Gives the key points of the text  
Bulletize: Formats the text into a bulleted list  
Elaborate: Adds details to build upon the text  
Shorten: Makes the text more concise  
Retry: Create a new version of the text.

After identifying places where a generated essay could be more concise or elaborated, for instance, students could prompt Gemini to elaborate on the ideas or shorten the text. Students could then critique the revised output and suggest further revisions. The integration of Gemini thus offers an opportunity to invite AI to critique its own outputs or for different AI tools to interact with one another, and for human and AI interlocutors to collaboratively compose, critique, or refine texts. By supporting students to probe into the layers of human- and AI-generated content, instructors could encourage students to critically interrogate generated feedback and to reflect on the human dimensions of composing.

Future iterations of this activity could encourage students to comment directly on one another's annotations using a social annotation tool such as Hypothesis or Perusall. Using the threaded discussions feature, students could respond to one another's annotations and build on or complicate others' ideas. By commenting on each others' contributions, posing questions, and interrogating assumptions, students could engage in a shared conversation about an AI-generated text while learning from their peers' contributions (Farber, 2019). In a follow-up reflection, students could trace the unfolding trajectory of a collaboratively annotated conversation and identify the ways the annotations shift or evolve over the course of a digital conversation. In addition, collaboratively annotating shared texts using Hypothesis or Perusall could stimulate students' engagements with multimodal texts such as images or web pages.

OpenAI's new Canvas feature for Plus users can further enhance the process of collaboratively composing with AI. By leveraging this feature, users can directly highlight and annotate specific portions of a text, prompting ChatGPT to revise a piece. In addition, users can receive feedback from ChatGPT by "highlight[ing] specific portions of your text or code and ask[ing] ChatGPT to either request edits to that section or ask a question about it" (OpenAI). The feature thus enables multidirectional interchanges between human and AI writers, facilitating dynamic,

iterative feedback processes. By leveraging the commenting tools embedded within the ChatGPT interface, students can engage in multiple, recursive, and instantaneous cycles of revision. In addition, by drawing on the “version history and show changes” feature students can reflect metacognitively on the writing process.<sup>6</sup> For instance, in a reflective writing exercise, students could examine how a piece of writing evolves over time: students might parse the ways certain writing choices have shaped a text or analyze the dialectical interweaving of human-AI interactions that unfold over the course of the revision process. Such an exercise could illuminate the fluid, malleable nature of writing as an evolving trajectory inflected by a constellation of intersecting choices. By immersing students in collaboratively composing with AI, instructors can inspire students to conceptualize writing as a layered, continually evolving process.

- Enhance students’ rhetorical awareness of writing across contexts.

A related possible recommendation would be to encourage students to question rule-based assumptions about writing, including grammar and essay structure, and to invite more rhetorical considerations of genre, audience, and purpose. For instance, students could prompt AI to compose in a variety of genres beyond a five-paragraph essay, such as reports, research papers, case studies, poems, or screenplays, and to examine the extent to which generated writing upholds or subverts genre conventions. Students could examine writing in particular disciplinary genres, with attention to the ways knowledge is constructed in particular disciplines and the extent to which generated writing exemplifies or falls short of disciplinary norms. In this sense, this activity could be adopted by other disciplines that may use writing, but not teach writing. Students could also comparatively analyze human- and AI-generated texts as a way to interrogate gaps or elisions in AI-generated writing and to attune to subtler distinctions between human and AI writing. Engaging in this annotation exercise can thus encourage students’ roles as co-constructors of knowledge who actively shape and critique artificially generated texts. Ultimately, this activity can open pathways toward innovative, accessible approaches to supporting students’ critical engagements with AI.

## References

- Alabed, A., Javornik, A., & D. Gregory-Smith (2022). AI anthropomorphism and its effect on users' self-congruence and self-AI integration: A theoretical framework and research agenda. *Technological Forecasting and Social Change*, 182, 1-19. <https://doi.org/10.1016/j.techfore.2022.121786>
- Al-khresheh, M. H. (2024). Bridging technology and pedagogy from a global lens: Teachers’ perspectives on integrating ChatGPT in English language teaching. *Computers and Education: Artificial Intelligence*, 6, 1-10. <https://doi.org/10.1016/j.caeai.2024.100218>.
- Anderson, S. S. (2023). Places to stand: Multiple metaphors for framing ChatGPT’s corpus. *Computers and Composition*, 68, 1-13. <https://doi.org/10.1016/j.compcom.2023.102778>.
- Barton, M., & K. Klint. (2011). A student’s guide to collaborative writing technologies. In C. Lowe & P. Zemliansky (Eds.), *Writing spaces: Readings on writing* (pp. 320-332). Parlor Press.
- Benjamin, W. (2008). *The work of art in the age of mechanical reproduction* (J. A. Underwood, Trans.). Penguin Books.

---

6. Chamberlain’s article “Draftback to the Future” offers an approach to encouraging students to reflect metacognitively on their writing processes by analyzing Google Docs revision histories.

- Brown, D. W., & Aull, L. L. (2017). Elaborated specificity versus emphatic generality: A corpus-based comparison of higher- and lower-scoring Advanced Placement exams in English. *Research in the Teaching of English*, 51(4), 394–417. <http://www.jstor.org/stable/44821273>.
- Bruffee, K. A. (1984). Collaborative learning and the conversation of mankind.” *College English*, 46(7), 635–652. <https://doi.org/10.2307/376924>.
- Chamberlain, E. (2018). Draftback to the future: A tool for writing process analysis. *Journal of Interactive Technology and Pedagogy*. <https://cuny.manifoldapp.org/read/draftback-to-the-future-a-tool-for-writing-process-analysis-a0a727a1-29c3-405c-bb1e-9fab441c74eb/section/31c98d54-889e-44ef-9030-e39a86a6a6a7>
- Chan, J. W. W., & J. W. C. Pow (2020). The role of social annotation in facilitating collaborative inquiry-based learning. *Computers & Education*, 147. <https://doi.org/10.1016/j.compedu.2019.103787>.
- Chatfield, T. (2013). I type, therefore, I am. AEON. <https://aeon.co/essays/the-world-is-awash-with-more-text-than-ever-before>
- Farber, M. (2019). Social annotation in the digital age. *Edutopia*. <https://www.edutopia.org/article/social-annotation-digital-age>.
- Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence*. Hillsdale, NJ: Erlbaum.
- Google. (2024). Write with Gemini in Google Docs. <https://support.google.com/docs/answer/13951448?hl=en>
- Kalir, J. H. (2020). Social annotation enabling collaboration for open learning. *Distance Education*, 41(2), 245–260. <https://doi.org/10.1080/01587919.2020.1757413>.
- Kim, S. H. (2015). Preparing English learners for effective peer review in the writers’ workshop. *The Reading Teacher*, 68(8), 599–603. <https://doi.org/10.1002/trtr.1358>
- Knoth, N. A. Tolzin, A. Janson, & J. M. Leimeister. (2024). AI literacy and its implications for prompt engineering strategies. *Computers and Education: Artificial Intelligence*, 6: 1-14. <https://doi.org/10.1016/j.caeai.2024.100225>.
- Li, L., Liu, X. and Steckelberg, A.L. (2010). Assessor or assessee: How student learning improves by giving and receiving peer feedback. *British Journal of Educational Technology*, 41: 525-536. <https://doi.org/10.1111/j.1467-8535.2009.00968.x>
- Li, R. (2024). A ‘dance of storytelling’: Dissonances between substance and style in collaborative storytelling with AI. *Computers and Composition* 71: 1-10. <https://doi.org/10.1016/j.compcom.2024.102825>
- MLA-CCCC Joint Task Force on Writing and AI Working Paper: Overview of the Issues, Statement of Principles, and Recommendations. (2023). MLA/CCCC. <https://aiandwriting.hcommons.org/working-paper-1/>.
- Markey, B., Brown, D. W., Laudenschach, M., & Kohler, A. (2024). Dense and Disconnected: Analyzing the Sedimented Style of ChatGPT-Generated Text at Scale. *Written Communication*, 0(0). <https://doi.org/10.1177/07410883241263528>
- Milroy, J., & Milroy, L. (2012). Authority in language: Investigating standard English. Routledge.
- Morales, E., Kalir, J. H., Fleerackers, A., & Alperin, J. P. (2022). Using social annotation to construct knowledge with others: A case study across undergraduate courses. *F1000Research*, 11, 235. <https://doi.org/10.12688/f1000research.109525.2>.
- Nagy, W. E., & Anderson, R. C. (1995). Metalinguistic awareness and literacy acquisition in different languages. *Center for the Study of Reading Technical Report; no. 618*.

- OpenAI. (2024). What is the canvas feature in ChatGPT and how do I use it? <https://help.openai.com/en/articles/9930697-what-is-the-canvas-feature-in-chatgpt-and-how-do-i-use-it>
- Roehr-Brackin, K. (2018). *Metalinguistic Awareness and Second Language Acquisition* (1st ed.). Routledge. <https://doi.org/10.4324/9781315661001>.
- Schleppegrell, M. J. (2013). The Role of Metalanguage in Supporting Academic Language Development. *Language Learning*, 63, 153-170. <https://doi-org.proxy.lib.umich.edu/10.1111/j.1467-9922.2012.00742.x>.
- Thieme, K. (2022). Spacious Grammar: Agency and Intention in the Teaching of Research Writing. *Discourse and Writing/Rédactologie*, 32, 281–299. <https://doi.org/10.31468/dwr.931>
- Tremain, L., Anderson, J., Eschenbach, B., Amann, N., & Marsden, K. (2023). What Counts as Literacy in the Polytechnic Hispanic Serving Institution? Culturally Sustaining Frameworks for Writing Assignments, Assessment, and Language Use. *Humboldt Journal of Social Relations*, 45, 119–135. <https://www.jstor.org/stable/48725123>
- Tsao & Nogues (2024). Beyond the author: Artificial intelligence, creative writing and intellectual emancipation, *Poetics*, 102, 1-12.
- Wigglesworth, G., & N. Storch (2012). What role for collaboration in writing and writing feedback. *Journal of Second Language Writing*, 21(4), 364-374. <https://doi.org/10.1016/j.jslw.2012.09.005>.
- Zhu, X., Chen, B., Avadhanam, R.M., Shui, H. and Zhang, R.Z. (2020). Reading and connecting: using social annotation in online classes. *Information and Learning Sciences*, Vol. 121 No. 5/6, 261-271. <https://doi.org/10.1108/ILS-04-2020-0117>.

## **Appendix A: Guidelines for Collaboratively Annotating ChatGPT-Generated Essays** (in-class activity)

### **Part I:**

In small groups (2-3 members each):

Annotate (add comments on) the ChatGPT-generated essay below.

Add 2-3 comments per paragraph.

You could consider questions such as:

- How specific, complex, or nuanced are the ideas in the essay? Where are places that could be more specific, complex, or nuanced? Which ideas could be developed further? Where could examples be added? Which alternative perspectives could be considered? To what extent does the essay present original or surprising ideas?

### **Title: Pros and Cons of Social Media: Navigating the Digital Landscape**

**Group 1:** Social media has become an integral part of modern life, offering numerous benefits alongside significant drawbacks. This essay examines the pros and cons of social media, highlighting its impact on individuals, communities, and society at large.

**Group 2:** On the positive side, social media facilitates connectivity and communication on a global scale. It allows individuals to stay in touch with friends and family, regardless of geographical distance, fostering meaningful relationships and support networks. Social media platforms also serve as powerful tools for information dissemination, enabling rapid sharing of news, ideas, and resources. Moreover, social media provides a platform for self-expression and creativity, empowering users to share their thoughts, talents, and experiences with a wide audience.

**Group 3:** However, the proliferation of social media also brings forth several challenges and drawbacks. One major concern is the prevalence of misinformation and fake news, which can spread rapidly and undermine public discourse and trust in credible sources. Social media platforms are also associated with issues like cyberbullying, harassment, and online trolling, which can have detrimental effects on individuals' mental health and well-being. Furthermore, excessive use of social media has been linked to addictive behaviors, leading to decreased productivity, sleep disturbances, and negative impacts on offline relationships.

**Group 4:** Privacy is another significant issue in the realm of social media, as users' personal data are often collected, analyzed, and monetized by platform providers and third-party advertisers. This raises concerns about data security, consent, and the ethical implications of targeted advertising and algorithmic manipulation. Additionally, the curated nature of social media feeds can contribute to echo chambers and filter bubbles, limiting exposure to diverse perspectives and fostering polarization in society.



**Group 5:** In conclusion, social media offers numerous advantages in terms of connectivity, information sharing, and self-expression. However, its drawbacks, including misinformation, cyberbullying, privacy concerns, and addictive behaviors, underscore the need for responsible usage and thoughtful regulation. By addressing these challenges and fostering a culture of digital literacy and ethical engagement, we can harness the potential of social media while mitigating its negative impacts on individuals and society.

## **Part II:**

In small groups (2-3 members each):

1. Prompt ChatGPT (or another tool of your choice) to write an essay (or a report, a case study, an abstract, etc.) on a topic of your choice.
2. Copy/paste the text below.
3. Annotate (add comments on) the AI-generated text. Add 2-3 comments per paragraph.
4. If you would like, prompt the tool to revise the text based on your comments or critiques.
5. Copy/paste the revised text below.
6. Annotate (add comments on) the revised AI-generated text.

**Appendix B: Guidelines for Individually Annotating ChatGPT-Generated Essays**  
(a short answer question included in the course midterm)

Read the AI-generated essay below.

**Navigating the Landscape of Artificial Intelligence: Affordances and Limitations**

Introduction:

Artificial Intelligence (AI) stands as a transformative force, reshaping industries, and augmenting human capabilities. Its advancements offer a spectrum of affordances and limitations that warrant examination. This essay delves into the multifaceted aspects of AI, exploring its potentials and constraints.

Affordances of AI:

AI's automation capabilities streamline repetitive tasks, boosting productivity and efficiency. Additionally, its prowess in data analysis unlocks valuable insights from vast datasets, enhancing decision-making processes. Moreover, AI enables personalized experiences by analyzing user data, fostering engagement and satisfaction. Its problem-solving prowess drives innovation, pushing technological boundaries. Lastly, AI enhances accessibility, offering assistive tools for individuals with disabilities, promoting inclusivity.

Limitations of AI:

Despite its potential, AI grapples with inherent biases, perpetuating societal inequalities. Its lack of creativity and contextual understanding limits its applicability in nuanced tasks. Moreover, AI's performance hinges on data quality, susceptible to errors and biases. Ethical and legal concerns loom large, from privacy issues to job displacement. Additionally, AI's vulnerability to security risks poses challenges for cybersecurity experts.

Conclusion:

Artificial Intelligence offers promising affordances across various domains, from automation to accessibility. However, its limitations, including bias and ethical concerns, demand careful consideration. As AI continues to evolve, it's crucial to navigate these challenges responsibly, ensuring it remains a tool for augmenting human intelligence. By addressing its limitations while harnessing its potentials, we can steer AI towards a future that benefits humanity.

1. Annotate (add comments on) places where the essay could be more specific (clearly defined or identified) and complex (considering various perspectives or aspects of an issue). Add 2-3 comments per paragraph.
2. Write a short paragraph (5-8 sentences) reflecting on what you noticed about the essay and identifying ways the essay could be revised to be more specific and complex.