

Critical Thinking and Artificial Intelligence: Grappling with ChatGPT in the Undergraduate Music Theory Classroom

HANNAH BENOIT

This article outlines the details of a critical thinking and Artificial Intelligence (AI) curriculum unit to give instructors one suggestion for how they might productively include the use of AI in their music theory classrooms. I share the different lessons and the assignment that I provided for my students, followed by a report of student responses and my assessment strategies. I conclude this article by encouraging music theory instructors to bring discussions of AI into the classroom in order to promote more meaningful learning outcomes—ones that extend beyond simply teaching students that ChatGPT and other Large Language Models (LLMs) are inappropriate to use. This AI unit encourages undergraduate music theory students to more thoroughly understand and experience the impacts of LLMs on critical-thinking skills.



Introduction

Artificial Intelligence (AI) is reshaping the way students think, write, and learn in higher education.¹ Some students turn to LLMs (Large Language Models) to help lighten their workload when formulating ideas and writing prose for a music-theoretical paper. Many of these students may not fully grasp the ramifications or potential pitfalls of relying on the software, especially when it comes to issues like academic integrity, authorship, ethical boundaries, or what is lost without their independent thought.² AI challenges traditional definitions of plagiarism and collusion, prompting institutions to reevaluate their academic honesty policies; however, many institutions lack clear guidelines because AI, especially LLMs, are still relatively new in educational settings.³ As a result, both students and instructors are often uncertain about what is acceptable. As the trend grows, it is becoming increasingly important to equip students with the skills to engage critically with AI to prepare them for the

1 See Du, X., Du, M., Zhou, and Bai (2025) and Walter (2024) for recent studies that evaluate how AI influences how students think and its impact on higher education.

2 See Shahzad et. al. (2025).

3 Viet (2025) offers a literature review that suggests the need for clear AI policies in higher education. Sullivan et. al. (2023, 4) reports that most universities either are undecided about their policies or outright ban ChatGPT.

professional world.⁴ Compounding the issue is that many instructors feel unequipped to meaningfully address AI in the classroom, often choosing to sidestep the topic entirely due to unfamiliarity with how these tools work or how to engage students in thoughtful discussions about their use. Without guidance, educators risk missing the opportunity to guide student development as reflective and responsible users of AI. In this paper, I encourage educators to cede some of their authority and engage in open, collaborative discussions about the implications and impact of LLMs, such as ChatGPT, in the music-theory classroom. I offer one way of doing this by including a short unit, based on a pilot study, focused on the issues and ethics of using these technologies in music theory and analysis.⁵

In summer 2024, I designed and taught an accelerated four-week section of Music Theory and Analysis IV to a small group of seven students, which included a one-week-long “Music Theory and AI” unit.⁶ This course, the fourth semester of a five-term music theory core sequence, focuses on the use of expanded harmonic resources of the nineteenth century (e.g., advanced chromaticism including enharmonic reinterpretation and symmetrical divisions of the octave), analysis of characteristic small and large forms, and writing and analytical skills, with a goal of students learning to articulate convincing analyses of selected repertoire. I structured the course into four one-week units, each focusing on either one type of Romantic-era musical technique, genre of repertoire (programmatic orchestral movements and character pieces), or formal archetype (song forms and ternary form) for analytical exploration. In addition to studying the musical parameters or genres that guided the content of each week, students were asked to complete three end-of-unit capstone projects, which included: 1. writing an analytical essay; 2. an assignment on AI and critical thinking; and 3. a recomposition assignment.⁷ I positioned the AI unit as the second of the four weeks, using programmatic orchestral music as the repertoire basis

4 Irish et. al. (2025) addresses the benefits of incorporating AI into the classroom to better equip students with AI knowledge applicable to workplaces outside of the university, and bolstering ethical awareness of these tools as well.

5 See Miserandino (2025) on supporting the incorporation of AI into the classroom and assignments for positive learning outcomes.

6 While this article presents the AI unit as compressed into a single week, lessons and assignments within this unit can be adapted for as little as one class session. The unit, as a whole, can be scaled to fit within a typical semester module comprising five to six class periods.

7 The last week culminated in my students’ final project, in which they expanded their analytical essay, recomposition assignment, or previously assigned in-class teaching demonstration.

for assignments and in-class discussions. The previous week required students to engage with and analyze Romantic-era song forms, culminating in an analytical essay in which they presented their interpretations of a musical work based on an in-class analysis using formal prose. Throughout a few class meetings, I shared models of what a thesis statement looks like, its function, and how to write one for an analytical paper. This essay assignment included my feedback and revisions, with a strong focus on their thesis statements and writing mechanics. The AI unit elaborates on the prior week's learning objective, which emphasized the development of critical-thinking and writing skills. I asked students to apply these critical-thinking and writing skills by assessing both themselves and ChatGPT.⁸

This article, outlines the details of the AI unit I developed, which asks students to articulate and support a judgement about the affordances and limitations that ChatGPT offers in the music-theory classroom, in order to present instructors with a model for how they might productively include LLMs—including ChatGPT, Gemini, DeepSeek, and others—in their classrooms. I share the lessons and assignment I provided for my students during this curriculum unit, followed by a report on student responses and my assessment strategies. The article concludes by encouraging music-theory instructors to bring discussions of AI into the classroom in order to promote more meaningful learning outcomes—ones that extend beyond simply teaching students that ChatGPT and other LLMs are inappropriate to use. While the potential for AI to inhibit critical-thinking skills warrants concern from instructors, this unit was designed to test that assumption by putting students in active, evaluative roles where they critically analyzed AI-generated writing. This pilot study encouraged undergraduate music-theory students to more thoroughly understand and experience the impacts of LLMs on critical-thinking skills.

A focus on critical-thinking skills is apt for this curriculum unit because AI poses a danger to developing these skills due to “cognitive offloading,” where individuals utilize external tools to reduce the mental effort required to complete the task at hand.⁹ If students rely too heavily on LLMs to “think” for them when completing assignments, then the result is stagnant knowledge or absorbing incorrect information as truth, which in turn will snowball to eventually impact the furthering of education and educational institutions’ core principles of critical thinking, intellectual growth, and

8 While any LLM can be used for this proposed unit, I selected ChatGPT due to its widespread popularity and the familiarity my students already had with the platform. At the time of this assignment, students used ChatGPT 3.5 Turbo, the free version available without subscription.

9 Gerlich (2025).

independent inquiry. In many higher-education classrooms, AI is not discussed openly, resulting in students who will continue to use ChatGPT and other LLMs while unaware of, or indifferent to, the greater impacts that these interfaces pose on their personal and collective educations.¹⁰ I propose this unit to help students better understand the environment they are currently navigating—to discuss and wrestle with the context in which ChatGPT currently exists in order to better educate the current generation of students. By the end of this unit, successful students are able to articulate a sound judgement of ChatGPT, how it interacts with their own learning and autonomy, and specifically the limitations of the interface’s music-analytical skills. I offer this unit to harness the power of AI towards something positive¹¹—instead of pretending that it does not exist or is too difficult to bring into the music-theory classroom, I encourage and show the constructive and powerful learning outcomes that can be produced by grappling with subjects, such as AI, that educators and students still have much to learn about.

The AI Unit

In order to create a comfortable classroom environment in which each student felt encouraged and supported to share their own ideas and experiences, I began the AI unit by facilitating an in-class discussion where I first shared my personal experiences with ChatGPT. I openly told my students that I had limited experience with ChatGPT because I had only used it for a handful of non-academic searches. I then asked my students if they had used ChatGPT before, whether they had used it for a music-theory course, and what sorts of limitations or strengths they had observed from their own or others’ interactions with ChatGPT. Some students stated that they often used ChatGPT to complete essays for all of their courses, while others admitted that they had never interacted with the interface before. The majority of the students who commonly used ChatGPT communicated that devising input questions that would elicit an appropriate response was the most challenging component. Students observed that ChatGPT falls short in providing intricate details, but excels at big-picture ideas and outlines of short essays.

¹⁰ From this point forward, I use “ChatGPT” as a stand-in for any LLM, as it was the model used for this unit. After running this unit in summer 2024, I believe instructors should also invite discussion of Gemini, DeepSeek, Claude, and other LLMs, since many students may have experience beyond ChatGPT.

¹¹ See Du, X., Du, M., Zhou, and Bai (2025) for AI’s role in enhancing higher-order thinking skills in students and Kshirsagar, et. al. (2022) for the encouraged use of AI in the classroom setting to leverage improved learning outcomes.

As part of this discussion, I also reviewed our university's academic integrity policy as it relates to AI tools, particularly around plagiarism, authorship, and ethical use.¹² This open-ended discussion allowed me to understand how students in a younger generation are interacting with LLMs, outside of situations in which I “catch” them using ChatGPT on an assignment in an inappropriate or dishonest way. I was also able to gauge how comfortable they would be using this tool for the upcoming end-of-unit “Critical Thinking and AI” assignment. Beginning this course unit with a plenary discussion not only promoted student comfort in sharing their ideas, but also situated the unit's learning objectives around the students' current understanding and knowledge of AI.

Building on these initial student perspectives, I realize that future iterations of this unit would benefit from including an introductory explanation of what LLMs are, including: how they differ from search engines, the sources from which they derive their information, and the resulting ethical and academic concerns their use raises. ChatGPT and other LLMs can reinforce various biases, which would encourage students to critically examine the social injustices these tools can perpetuate. This framing would help students better understand not only the capabilities of generative AI, but also its potential limitations—particularly with regard to critical thinking, academic integrity, and discrimination.¹³ Ethically, these tools raise concerns around the use of copyrighted materials in their training without permission. At the time of designing and teaching this unit, I was not yet aware of the negative ecological impacts of ChatGPT. These tools have high energy demands and contribute significantly to carbon emissions. I highly encourage facilitating a conversation about both ethical and ecological implications of generative AI with your students as part of the introduction to this course segment.¹⁴

In order to scaffold this unit and provide my students with formative feedback before taking on the “Critical Thinking and AI” project, the next stage of this curriculum unit was an in-class workshop with ChatGPT. I broke the class into small groups to

12 McGill University's policy on AI categorizes offences into 1. Low-to-moderate risk: using AI as a tool, which includes proofreading or using it as a search engine; 2. Moderate-to-high risk: using AI to generate content for one's use, encouraged as an option for learning and not for publication or sharing with others; and 3. Very high risk: using AI to generate content and presenting it as one's own work, which results in a violation of my university's academic policy.

13 Resources such as www.thebullshitmachines.com and <https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/> can provide accessible and provocative entry points for students to explore the persuasive and often misleading outputs of generative AI tools.

14 See <https://www.nature.com/articles/s41893-025-01681-y> for more information.

prompt ChatGPT to generate a short analytical essay based on “Meeres Stille” by Franz Schubert, which we had previously analyzed together in class. I observed and commented on each group’s interactions with ChatGPT, often encouraging each group to assess how their prompts were phrased and how they could manipulate subsequent input questions in order to elicit a response with which they were satisfied. I intentionally grouped students that have never used ChatGPT with the students that had interacted with the interface in order to promote peer learning and insightful conversations. Students exercised their critical-thinking skills by creating multiple input questions and refining them based on the outputs they received. This process asked students to evaluate and modify prompts when the responses deviated from the intended objective. This in-class workshop phase of the AI unit allowed my students to discover for themselves the limitations and possibilities of ChatGPT in terms of music analysis, and they were free to do so unencumbered by concerns about a grade. Students responded to this activity positively and commented that an instructor had never asked them to use ChatGPT during class time and in such a transparent manner. In future iterations of this unit, I will teach students common prompting strategies either before or after this workshop activity to scaffold the upcoming end-of-unit assignment.¹⁵

After this hands-on activity, I then shared the learning objectives and tasks for the end-of-unit project, which can be found in Appendix 1. This document outlines all components of this assignment’s learning objectives. My students received this assignment sheet prior to analyzing the *Tristan und Isolde* Prelude opening in order to direct and guide their experience of the piece in service of the final project. They focused on the hallmark characteristics of the piece, including Leitmotifs, formal structure, and tonal ambiguity. Students explored the Tristan chord in particular, discussing its many interpretations and its significance.¹⁶

The learning objective for the “Critical Thinking and AI” assignment is not only for students to demonstrate the knowledge gained about 19th-century chromatic harmony and techniques from our in-class analysis of the Tristan Prelude, but also to assess and question LLMs and their usages in the music-theory classroom and analytical writing. This assignment comprises three phases that required critical-thinking skills in multiple modalities. In phase 1, students prompted ChatGPT to provide a short essay discussing the aspects of the Prelude that we had talked about in class. Students

15 Resources such as Wang and Jin (2023, 5-7) or blog posts (<https://www.coursera.org/articles/how-to-write-chatgpt-prompts>) could bolster student engagement.

16 See Martin (2008, 6-30) for a survey of the Tristan chord’s various analyses.

exercised their critical-thinking skills in order to produce an input question that would elicit an output from ChatGPT that could go undetected by a professor. While I did not ask students to provide a list of their input questions, students stated in classroom discussion that ChatGPT only produced essays they could use if the prompt question was specific and provided enough context. One student provided their prompt in their submission:¹⁷

You are a fourth-year amateur music theory student in university with limited essay-writing skills. You need to write a 400-word essay about Wagner's Prelude to *Tristan und Isolde*, specifically focusing on the Tristan chord. The essay needs to include a creative and fitting title, a thesis statement that begins with "In this essay, I argue..." at the end of the first paragraph, three different musical examples with in-depth music theory related analysis. Do not use first-person pronouns anywhere other than in the thesis statement, do not use complicated words. First paragraph needs to start with an introduction, followed by the thesis statement.

This student was specific in positioning the essay's tone (an "amateur music theory student"), and specifically asked for the exact requirements outlined in the grading rubric.¹⁸ Students reported that the ChatGPT essays submitted for the assignment were not a result of their very first input questions.¹⁹ Instead, each student needed to assess ChatGPT's responses and subsequently change their prompt based on the output in order to provide a better essay that discussed specific points on the assignment sheet.

Phase 2 of this assignment exercised students' critical-thinking skills in a different modality, this time by correcting the ChatGPT essay. This course emphasizes writing skills and, more specifically, music-analytical writing skills. The rubric that my students used to assess ChatGPT's writing (See Appendix 2) is the same rubric I used to assess their previous end-of-unit analytical essay.²⁰ While I acknowledge that

¹⁷ Student work has been anonymized and consent to share quotes of their work was received from each student in this class.

¹⁸ Poor prompts are usually not this thorough, using one-sentence inputs such as "Write a 400-word essay about the Tristan Chord." The more vague and non-specific the input, the more vague and non-specific the output.

¹⁹ In our discussion, students primarily described refining their prompts to achieve an acceptable essay, but it remains unclear whether they also directly revised or rewrote ChatGPT's output once they received a usable draft. This would be a productive question to explore, as many users often shift from prompting to manual editing, and it would shed further light on the types of critical-thinking strategies students employ when working with LLM-generated text.

²⁰ Scholars such as McGee (1993), Attas (2016), Ripley (2020), and Rogers (2017) demonstrate that prose-based assignments—especially those that ask student to articulate, defend, or question analytical interpretations—can foster deeper engagement and independent thinking. See McGee (1993); Attas (2016); Ripley (2020); Rogers (2017) for discussions of critical-thinking skills as they relate to writing skills in the music-theory classroom, who's contribution influence the assessment focus of this rubric.

rubric elements such as “visual examples” may seem less relevant when assessing ChatGPT-generated work, I retained the original essay rubric for consistency with earlier student assessments. That said, the “visual examples” category is partially applicable in this context: while ChatGPT cannot currently generate musical notation, it can outline notes of chords and identify specific measure numbers. Future iterations of this assignment would likely revise this rubric element to better reflect the capabilities and limitations of generative AI tools—unless future developments in ChatGPT’s programming enable the interface to generate musical notation. For this assignment, I required students to import the essay into a Word document and use “track changes” to copyedit it and add comments to the text. Students were instructed to turn in the ChatGPT essay with these markups included so I could clearly identify what aspects of the essay they assessed. My assignment sheet provided a few starting points and required that students each make at least three revisions or comments on the essay. Communicating revisions and comments directly on the essay demonstrated students’ critical-thinking skills by asking them to assess and question what they were reading. The rubric guided the students’ revisions and comments and directed them to articulate a sound judgement about the essay and a numerical grade.

Phase 3 asked students to write a reflection about their experiences using AI tools, with the goal of articulating their attitudes and opinions regarding the place that ChatGPT may or may not have in music analysis.²¹ In these short reflections, I required each student to discuss the connection between critical thinking and AI. I asked students to provide an honest response to their interaction with AI tools and specifically address whether ChatGPT helped or hindered their personal critical-thinking skills. I also asked students to reflect on their experience; changes, if any, in their perception of AI tools; their opinion of AI in music-analytical writing; and anything else that they found valuable during their learning process. This emphasis on self-reflective practice supports learner agency and hones students’ abilities to critique systems of knowledge production, which includes technologies like ChatGPT, all in the service of better understanding the broader social and educational contexts in which they are learning. Student submitted the assignment, including the marked-up ChatGPT essay, the graded analytical essay rubric, and their short reflection.

21 See Ferenc (2016) for a discussion regarding how incorporating reflection as part of classroom activities can yield deeper and more self-regulated learning.

Student Responses and Assessment Techniques

I assessed student submissions using the rubric shown in Appendix 3.²² The rubric evaluated each student on their demonstrated ability to comment on and revise the LLM-generated essay, the content and writing mechanics used in their reflective response, and how accurately and thoughtfully they filled out the essay-grading rubric. The completion of this assignment made up 15% of my students' total course grade. This section will share examples of comments students wrote on the AI-generated essay, quotes from their reflective responses, and examples of their graded rubrics. In future iterations of this unit, I would require students to provide their dialogue with the LLM. Providing a copy of these interactions would allow me to observe each student's critical-thinking process as they assessed the LLM's outputs and changed their input based on these observations.

The essay mark-up portion of this assignment demonstrated students' command of knowledge about music-theoretical writing gained from the course's previous essay assignment. Reading my students' corrections on LLM-generated essays provided even deeper insight into their improved music-analytical writing skills. Example 1 shows some sample student corrections and comments on LLM-generated thesis statements to exemplify students' error-detection and critical-thinking skills. The student comments regarding LLM-generated thesis statements were much more thorough than I anticipated. These comments demonstrate that not only were students able to accurately identify the thesis statement of the AI-generated essay, but they were also able to critique the writing and suggest changes to the thesis statement for improved clarity. This error-detection exercise led to my desired learning outcomes: it fostered critical thinking, reinforced their understanding of thesis statements and music-analytical writing, and empowered them to take ownership of their knowledge by making corrections.²³ Critiquing the LLM complemented more traditional peer-review methods by offering students a low-stakes, external object of critique. Students practiced skills similar to those used in peer review (evaluating structure, clarity, and argumentation) without the interpersonal pressure of critiquing a classmate's work. By asking my students to take on the role of the teacher and comment on written text—just as we often do as instructors—I aim to not only deepen their critical-

22 My class and I created this rubric collaboratively, which facilitated a sense of trust, understanding, and agency in their assessment.

23 For studies regarding the affordances of error detection in prose writing see: Winograd and Johnston (1982, 61–76); Zamora et. al. (2018, 175–85) and Henshaw (2023, 31–41).

thinking skills and engagement with the material, but also to disrupt traditional power dynamics. By “disrupting traditional power dynamics,” I do not mean upending the classroom hierarchy entirely, but rather creating opportunities for students to step into evaluative roles, helping them see themselves as contributors to knowledge rather than passive recipients.

ChatGPT: I will analyze how Wagner’s deliberate avoidance of harmonic resolution and embrace of tonal ambiguity express the inner turmoil and desires of the characters in “Tristan und Isolde,” offering insight into the human condition through music.

Student comment: Thesis statement is good, but perhaps could focus more on the music analysis side than the insight into the human condition aspect. For example, “I will analyze...by...” – explaining how you will prove that the avoidance of harmonic resolution and tonal ambiguity work in this way.

ChatGPT: One of the most striking elements of the prelude is the use of the so-called Tristan Chord, a harmonic innovation that has been widely studied and debated for its profound impact on the subsequent development of music.

Student comment: Your thesis can be more specific, I’m not sure what you want to talk about. Maybe “One of the most striking elements of the prelude is the use of the Tristan Chord. In this essay, I argue the Tristan Chord is a four-note chord notable for its ambiguous and unresolved quality.” However, I’m not sure what exactly you are trying to analyze, and this thesis doesn’t seem to be very connected to the entire essay.

Example 1

Two thesis statements from ChatGPT-generated essays with student corrections and comments.

Example 2 illustrates student evaluations of other aspects of the ChatGPT-generated essay. Student comments range from identifying specific inaccuracies of musical analysis, to prodding ChatGPT to develop a more precise interpretation, to highlighting specific sentences that the student believed did not sound human-generated. The majority of my students submitted both types of comments: higher-order concerns about LLM-generated thesis statements and lower-level corrections of analytical details. The comments provided on the LLM-generated essays alone achieved two of my main learning objectives for this assignment: students were able to demonstrate their knowledge of 19th-century chromatic harmony and compositional techniques while also questioning, assessing, and correcting AI-generated music-analytical writing.

ChatGPT: Additionally, throughout the Prelude, Wagner employs a technique known as harmonic suspension, where unresolved chords linger, prolonging the sense of instability.

Student comment: A little unclear why the suspension and lingering is significant. Why does this create a sense of instability?

ChatGPT: This interpretation aligns with traditional functional harmony where secondary dominants create temporary tonicization, enriching the harmonic palette.

Student comment: Descriptive, but a human would never say that.

ChatGPT: The dissonance between the notes, particularly the unresolved nature of the augmented fourth (F to B), creates a feeling of unease and anticipation, setting the stage for the emotional drama that unfolds.

Student comment: This is a good sentence that describes how the chord manipulates with one's emotions, however, it lacks the theoretical analysis of the chord and the tonal significance of other augmented intervals present in the chord.

Example 2

Three comments from my students on ChatGPT's essays.

In addition to correcting the ChatGPT-generated essays, students critically evaluated the LLM's technical music writing by thoughtfully completing the same rubric that was used to assess their work on a previous assignment (shown in Appendix 3). Example 3 shows a complete ChatGPT-generated essay that one of my students commented upon and assessed; the generated content presented here is representative of what most students submitted. Example 4 includes excerpts from several students' rubric submissions, illustrating how they addressed five different sections of the rubric. Students meticulously assessed ChatGPT's analytical observations, the integrity and scope of its thesis statements, flow and logical ordering of ideas, polished prose, and ChatGPT's lack of supporting evidence and details by exercising their error-detection skills. The students' commentary on the various aspects of the AI-generated essay highlighted not only their growing ability to detect errors, but also their successful progress toward achieving my learning objectives. Students showed an understanding of what constitutes a strong analytical observation, recognizing that while LLMs may reference specific measures, its analytical details are often inaccurate, and it lacks musical examples to support descriptive analyses. Moreover, beyond music-theory-specific learning objectives and content acquisition, students conveyed their perceptions of what constituted strong writing skills. The rubrics students filled in led

them to articulate what a successful thesis statement requires, in addition to assessing writing mechanics such as tone, structure, and grammar.

"Harmonic Ambiguity and Emotional Depth: Analyzing Wagner's Prelude to 'Tristan und Isolde'"

In Richard Wagner's Prelude to "Tristan und Isolde," the composer embarks on a radical departure from traditional tonal structures, crafting a musical landscape that defies resolution and embraces ambiguity. Through the strategic use of chromaticism, harmonic suspension, and thematic transformation, Wagner immerses the listener in a world of profound emotional depth and complexity. I will analyze how Wagner's deliberate avoidance of harmonic resolution and embrace of tonal ambiguity express the inner turmoil and desires of the characters in "Tristan und Isolde," offering insight into the human condition through music.

Wagner's Prelude to "Tristan und Isolde" exemplifies his departure from traditional tonal structures through several distinct musical techniques. One notable example is the use of chromaticism, where the music shifts seamlessly between keys without clear resolution. For instance, the opening Tristan chord, consisting of F-B-D#-G#, defies conventional tonal expectations by containing dissonant intervals and lacking a clear tonal center. This chord establishes an atmosphere of tension and ambiguity from the outset, setting the stage for the emotional depth of the opera (approx. measures 1-3).

Additionally, throughout the Prelude, Wagner employs a technique known as harmonic suspension, where unresolved chords linger, prolonging the sense of instability. For example, the passage following the Tristan chord features unresolved dissonant seventh chords, such as B-D-F#-A b, which create a feeling of anticipation without providing the expected resolution to a tonic chord (approx. measures 4-8). This deliberate avoidance of harmonic resolution contributes to the overall sense of unease and yearning that pervades the music.

Furthermore, Wagner's use of thematic transformation adds to the tonal ambiguity of the Prelude. Leitmotifs associated with specific characters or emotions undergo constant variation and development, blurring the distinction between different tonal centers. For instance, the "Longing" motif, initially introduced in the cellos and basses, undergoes a series of transformations throughout the Prelude, moving through various keys and harmonic contexts (approx. measures 20-30). This fluidity of tonal identity reflects the complex emotional journey of the characters, as their desires and conflicts evolve over the course of the opera.

Another compelling example of thematic transformation in Wagner's Prelude to "Tristan und Isolde" is the evolution of the "Love" motif. Initially introduced with a tender and yearning quality, this motif undergoes a series of metamorphoses, evolving into more dynamic and expansive variations as the Prelude unfolds, with changes in orchestration, rhythm, and harmonic context. (approx. measures 40-50). From its gentle introduction in the strings to its grandiose restatement in the brass, the "Love" motif traverses a diverse musical landscape, embodying the multifaceted nature of love itself. This fluidity of expression mirrors the tumultuous relationship between Tristan and Isolde, as their love transcends societal norms and undergoes profound transformations throughout the opera.

In sum, Wagner's Prelude to "Tristan und Isolde" employs a range of musical techniques to create a sense of tonal ambiguity and uncertainty. From chromaticism and harmonic suspension to thematic transformation, the score is replete with examples of Wagner's innovative approach to tonality, plunging the listener into a world of profound emotional depth and complexity.

Example 3

One example of the ChatGPT-generated prose students were asked to evaluate.

Analytical observations (13/25): Analytical observations are musically insightful and accurate. All interpretive claims are clearly and convincingly linked to analytical observations.	You need to reference specific measure numbers in a specific score and place parenthetical measure numbers next to the chords/harmonies described. Analytical observations are accurate but need more concrete musical examples linked to the descriptive analysis.
Thesis Statement (4/10): Interpretation or “analytical tale” is clearly and specifically communicated to the reader through a thesis statement. The thesis begins with the words “In this essay, I argue...”	There is somewhat of a thesis statement that is not clearly stated at all. It was just randomly thrown into the introduction with no further explanation or summarized evidence supporting it. The paper’s main aim is to discuss the tonal ambiguity of the Prelude, mainly focusing on the Tristan Chord. I made a comment on the word document claiming that I am assuming these is a thesis statement (even I am unsure if it is ever claiming a main focus). It was not organized, and definitely rushed to get into the “nitty gritty” of the paper, instead of taking care of the foundations of an argumentative paper.
Flow and logical ordering of ideas (15/25): Ideas are presented in a logical order and arguments support the thesis of the project; a creative “analytical tale” is presented and all observations contribute to the narrative.	This essay was well constructed, that is to say that it has an introductory sentence, a thesis, paragraphs, and a conclusion. However, it is lacking flow in a significant way. The essay mentions different instances where the Tristan chord appears but each time the interpretation is more or less the same. The ideas in each paragraph does not flow in a way that is progressively more convincing as the essay goes on. These is also not a “tale” happening, the writer just repeats similar analysis in different wording throughout.
Visual examples (7/20): Examples are neat and tidy; different example formats are used to show different aspects of the analysis; examples are integrated at appropriate times throughout and clearly demonstrate the analytical point being made.	Points are deducted for the lack of visual examples (ChatGPT is unable to incorporate visual examples to accompany its arguments, for example annotated scores). However, 7 points are awarded for its effort to use a several examples to back up its ideas. For example, “[The Tristan Chord] makes its striking entrance early in the piece, notably at measure 3, immediately setting a tone of profound ambiguity.” At least the reader knows where to look as it clearly states that the first appearance of the chord is at measure 3.
Polished prose (10/15): Project meets minimum length requirements; clear attention to proofreading and editing results in no grammatical or spelling errors; sentence structure is flawless.	The essay met the word count. Minimal grammatical errors, however some parts were wordy and certain sentence structures did not make sense.

Example 4

Five rubric comments filled out by students.

Student responses to the LLM-generated essay, accompanied by their completed rubrics, highlight their developing understanding of 19th-century chromatic harmony and compositional techniques, while also showcasing students’ abilities to critically evaluate and articulate the limitations of AI-generated content. Students also provided

reflective responses that extended beyond music-theory disciplinary boundaries, connecting to real-world learning objectives such as transferable analytical and evaluative skills. For the last component of this assignment, students were asked to reflect upon how AI interacts with critical-thinking skills, which provided space for students to share their opinions about ChatGPT, their attitudes towards what place AI tools might have in the music theory classroom, and general commentary about what they learned through completing this assignment. Example 5 shares excerpts from the reflective essays, exemplifying how students connected AI use to contexts beyond the music-theory classroom and articulating their biggest take-aways from this assignment.²⁴

The primary consensus among my students was that while LLMs are a powerful tool capable of reducing their workload, its use is not without significant trade-offs. Students observed that using AI is not only detrimental to their own individual learning paths, but also poses a significant threat to collective human knowledge. Students recognized the danger AI poses towards music-analytical writing, and also communicated their conclusions about how AI cannot perform in the same way as humans are able to, especially when it comes to recognizing specific musical details and creating analytical interpretations.

²⁴ Student comments use the terms AI and ChatGPT interchangeably, a common conflation of the broader term AI with the specific LLM ChatGPT.

“The connection between critical thinking and AI became apparent as I began; I can say that, overall, I experienced much less demand on my critical thinking skills to complete the project than if I sat down with the music score, read some scholarly papers, and pieced together my own conclusions.”

“In the end, it was a useful assignment. There will always be students who rely on AI but learning how to interact with AI instead of ignoring it provides value.”

“Throughout the exercise, I realized that AI struggles to deeply analyze music, often relying heavily on existing internet content rather than discovering new ideas from the music itself.”

“I think AI takes a lot away from students such as critical thinking, evaluating sources, and writing skills because it packages everything so nicely without citing a single source or giving any credit of where it got its information.”

“GPT will make humans stop thinking.”

“If I was a teacher grading this, I would have recognized its AI. The essay lacked structure, an argument, and there was little to no support. There was also some odd wording that did not sound too authentic.”

“AI cannot critically assess its arguments’ strength...an AI generated essay gives off a stagnant flow of ideas.”

Example 5

Quotes from students’ reflective essays after completing the assignment.

Conclusion

This article presents a curriculum unit focused on AI, presented as a starting point for any music-theory instructor who is concerned that AI tools hinder students’ learning. I initially shared similar concerns about the impact of AI on students’ independent thinking. However, students’ thoughtful engagement with AI output in this unit challenged that assumption. I encourage educators to bring conversations about AI into the open in order to start a dialogue with students about the dangers it poses to their knowledge. This unit encourages students to experience firsthand how ChatGPT performs when asked to produce music analysis. While the unit outlined above was structured for one week of an accelerated course, my aim is to provide music-theory educators with a starting point to integrate AI for even one class to spark these important conversations with students.

The main learning objectives of this unit include challenging students to demonstrate their knowledge of 19th-century compositional techniques and exercise their critical-thinking skills through their assessment of ChatGPT. In addition,

I encouraged students to consider the affordances and limitations of AI usage, specifically urging them to articulate how it directly impacts their own critical thinking. The results of this unit were very positive overall. Not only did the students' reflective responses communicate that they had achieved outlined learning objectives, but many also expressed appreciation for the treatment of ChatGPT as a topic for dialogue rather than an academic offense.

Perhaps most compelling was the shift in students' attitudes towards ChatGPT. They began to see it not just as a shortcut, but as something to question, evaluate, and learn from. There is no guaranteed way to stop students from using AI tools, and as the technology progresses, I know that it will become harder and harder for educators to recognize what is generated by their students and what is generated using AI tools. Instead, I assert that a way to combat what could become an epidemic of AI-related loss of critical-thinking skills is for educators to question and challenge AI within the curriculum. Students using LLMs to complete essays cannot be overlooked—critical-thinking and writing skills are at stake.²⁵ We, as music-theory educators, must raise the bar in what we ask students to perform in order to challenge this threat to higher education. My small group of students were able to articulate sound judgements and opinions about ChatGPT and its interaction with music-analytical writing. My aim was to grant them the knowledge and ability to understand what choices they make when they turn to a large language model to aid them in completing an assignment. This assignment encourages students to critically examine, question, and assess ChatGPT's role in the music theory classroom. The unit empowers students to continue to question AI usage, all while disrupting traditional power dynamics by treating the instructor as an equal participant in our collective exploration of AI tools.

25 For a broader articulation of this concern within higher education, see a recent Chronicle of Higher Education reflection by a faculty member serving on a university AI committee, who observes that debates about AI increasingly raise the unsettling question, "Do [students] still need to learn how to write a paper?" The article suggests that preservation of writing, and the critical thinking it cultivates, remains essential. Megan Fritts, "What I Learned Serving on My University's AI Committee," *The Chronicle of Higher Education*, May 23, 2025. <https://www.chronicle.com/article/what-i-learned-serving-on-my-universitys-ai-committee>.

Works Cited

- Attas, Robin. 2016. "Teaching What We Do and How We Do It: Using a Miniconference Assignment to Dig Deep into Musical Analysis." *Journal of Music Theory Pedagogy* 30.
- Du, X., Du, M., Zhou, Z. et al. 2025. "Facilitator or Hindrance? The Impact of AI on University Students' Higher-Order Thinking Skills in Complex Problem Solving." *International Journal of Education Technology in Higher Education* 22, no. 39. <https://doi.org/10.1186/s41239-025-00534-0>.
- Ferenc, Anna. 2016. "Promoting Metacognitive Reflection in Music Theory Instruction." *Journal of Music Theory Pedagogy* 30, no. 2. <https://doi.org/10.71156/2994-7073.1183>.
- Gerlich, Michael. 2025. "AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking." *Societies* 15, no. 1: 6. <https://doi.org/10.3390/soc15010006>.
- Henshaw, Florencia. 2023. "Catch My Errors If You Can: The Relative Value of Peer Editing in the Language Classroom." *Northeast Conference on the Teaching of Foreign Languages*, no. 90: 31–41.
- Irish, Andrea L., Michele W. Gazica, and Vincent Becerra. 2025. "A Qualitative Descriptive Analysis on Generative Artificial Intelligence: Bridging the Gap in Pedagogy to Prepare Students for the Workplace." *Discover Education* 4, no. 48. <https://doi.org/10.1007/s44217-025-00435-4>.
- Kshirsagar, Pravin R., D. B. V. Jagannadham, Hamed Alqahtani, et al. 2022. "Human Intelligence Analysis through Perception of AI in Teaching and Learning." *Computational Intelligence and Neuroscience* 2022, no. 1. <https://doi.org/10.1155/2022/9160727>.
- Martin, Nathan. 2008. "The Tristan Chord Resolved." *Intersections: Canadian Journal of Music* 28, no. 2.
- McGee, Deron L. 1993. "The Power of Prose: Writing in the Undergraduate Music Theory Curriculum." *Journal of Music Theory Pedagogy* 7.
- Miserandino, Marianne. 2025. "Authentic and Creative Assessment in a World with AI." *Teaching of Psychology* 52, no. 3: 267–72. <https://doi.org/10.1177/00986283241260370>.
- Ripley, Angela. 2020. "Post-Tonal Postcards: Communicating Analysis and Reflection through Prose Writing." *Journal of Music Theory Pedagogy* 34.
- Rogers, Lynne. 2017. "Asking Good Questions: A Way into Analysis and the Analytical Essay." *Journal of Music Theory Pedagogy* 31.
- Shahzad, Muhammad Farrukh, Shuo Xu, and Hira Zahid. 2025. "Exploring the Impact of Generative AI-Based Technologies on Learning Performance through Self-Efficacy, Fairness & Ethics, Creativity, and Trust in Higher Education." *Education and Information Technologies* 30, no. 3: 3691–3716. <https://doi.org/10.1007/s10639-024-12949-9>.
- Sullivan, Miriam, Andrew Kelly, and Paul McLaughlan. 2023. "ChatGPT in Higher Education: Considerations for Academic Integrity and Student Learning." *Journal of Applied Learning & Teaching* 6, no. 1. <https://doi.org/10.37074/jalt.2023.6.1.17>.
- Walter, Yoshija. 2024. "Embracing the Future of Artificial Intelligence in the Classroom: The Relevance of AI Literacy, Prompt Engineering, and Critical Thinking in Modern Education." *International Journal of Educational Technology in Higher Education* 21, no. 15. <https://doi.org/10.1186/s41239-024-00448-3>.

- Wang, Shuyue, and Pan Jin. 2023. "A Brief Summary of Prompting in Using GPT Models." *Qeios*. <https://doi.org/10.32388/IMZl2Q>.
- Winograd, Peter, and Peter Johnston. 1982. "Comprehension Monitoring and the Error Detection Paradigm." *Journal of Reading Behavior* 14, no. 1: 61-76. <https://doi.org/10.1080/10862968209547435>.
- Viet, Nguyen Khoa. 2025. "The Use of Generative AI Tools in Higher Education: Ethical and Pedagogical Principles." *Journal of Academic Ethics*. <https://doi.org/10.1007/s10805-025-09607-1>.
- Zamora, Ángela, Suárez, José Manuel, and Diego and Ardura. 2018. "Error Detection and Self-Assessment as Mechanisms to Promote Self-Regulation of Learning among Secondary Education Students." *The Journal of Educational Research* 111, no. 2: 175-85. <https://doi.org/10.1080/00220671.2016.1225657>.

Appendices: Resources for Critical Thinking and AI Unit

Critical Thinking and AI

Learning Objective: Students will demonstrate their knowledge of 19th century chromatic harmony and compositional techniques used in programmatic orchestra music. Through writing a short reflective essay, they will question, assess, and correct AI-generated music analysis.

Without warning, your theory instructor has assigned you a short essay (300-500 words) to summarize the analytical interpretations of the Tristan Prelude that you've discussed in class—due tomorrow! You already have to practice tonight, feed your neighbor's cat, run a half marathon, and go grocery shopping.... there is no way for you to fit this all in, so you turn to ChatGPT to help you out in your time of need. You hope that the computer-generated short essay will be 1. Good enough to help you pass the assignment and 2. Go undetected as a document not written by you.

Phase 1

Step 1: Pick one or two of the following aspects of the “Prelude to Tristan und Isolde” discussed in class:

- Leitmotifs
- Tonal ambiguity
- Formal structure
- Tristan Chord

Step 2: Ask ChatGPT (or other LLM of your choice—please specify what you use) to write you a 300–500 word essay summarizing your chosen musical aspects. Play around with your input question—pretend like you are genuinely trying to trick me into thinking you wrote this summary and response! (A question to get you started: “Write me a 300–500 word summary about Wagner’s Prelude to Tristan und Isolde, specifically focusing on leitmotifs”).

Step 3: Once you are happy with ChatGPT’s response, copy and paste the response into a Word document.

Phase 2

Step 4: You are now the teacher! Correct ChatGPT’s response by either printing out the document and marking it up or adding comments and making edits using “track changes” in the Word document itself. Please provide a minimum of 3 corrections of grammar, structure, credibility of analytical statement, or just weird wording in general. Ask yourself:

- Is there a thesis statement?
- Does this sentence sound like AI and why? Would I have written this myself?
- Is this wording odd or vague?

Step 5: Fill out the “Analytical Essay Rubric” to give ChatGPT a grade.

Phase 3

Step 6: Provide a short (200-300 word) reflection about where you think ChatGPT excelled or really fell short in writing a music analytical essay. You must include in your response a reflection on the connection between critical thinking and AI (i.e. what does AI take away from students and learning?) and one or

more of the following points:

- Your own opinions of using AI.
- Whether this assignment changed your thoughts about using ChatGPT.
- What place you think ChatGPT might have in the music theory classroom.
- If you were the teacher, would you have recognized this as AI generated?

Please submit three documents for this assignment:

- The ChatGPT essay with comments/corrections.
- The graded "Analytical Tale Rubric."
- Your short reflection response.

Criteria <i>Standards for this performance</i>	Feedback
Title (/5): The title of the project is creative and gives a clear and nuanced idea of what the paper is about.	
Thesis statement (/10): Interpretation or “analytical tale” is clearly and specifically communicated to the reader through a thesis statement. The thesis begins with the words “In this essay, I argue...”	
Polished prose (/15): Project meets minimum length requirements; clear attention to proofreading and editing results in no grammatical or spelling errors; sentence structure is flawless.	
Visual examples (/20): Examples are neat and tidy; different example formats are used to show different aspects of the analysis; examples are integrated at appropriate times throughout and clearly demonstrate the analytical point being made. ²⁶	
Analytical observations (/25): Analytical observations are musically insightful and accurate. All interpretive claims are clearly and convincingly linked to analytical observations.	
Flow and logical ordering of ideas (/25): Ideas are presented in a logical order and arguments support the thesis of the project; a creative “analytical tale” is presented and all observations contribute to the narrative.	

Appendix 2

Analytical essay rubric students used to assess ChatGPT’s essay.²⁷

²⁶ At this time, ChatGPT cannot generate musical notation. This section of the rubric was included in students’ original grading rubric for their analytical essays, in order to keep assessment consistency. I would recommend removing the portion of the grading rubric unless newer versions of ChatGPT may include the ability to include musical notation.

²⁷ Analytical essay assignment and rubric created by Dr. Alexandria Jonker (State University of New York at Potsdam).

Criteria <i>Standards for this performance</i>	Feedback
AI Essay mark-ups (/20): The AI generated essay is clearly marked up, with a minimum of three comments regarding, analytical credibility, awkward wording, missing thesis statement, or vagueness.	
Content in Reflective Response (/30): The reflection demonstrates engagement with and understanding of this assignment through commentary about the connection between critical thinking and AI, in addition to a reflection about your opinion about using this tool.	
Grammar and Writing Techniques in Reflective Response (/25): The reflection contains good grammar, structure, and flow.	
Graded “Analytical Tale Rubric” (/25): The rubric is completely filled out with thoughtful comments and an appropriate grade reflecting ChatGPT’s performance.	

Appendix 3

Grading rubric for “Critical Thinking and AI” assignment.