

RAPID: DRL AI: The Development of a Digital Platform for Evaluating and Using AI-Generated Content for Academic Purposes (NSF Award # 2337969)

Dr. Amy Hutchison (PI), Dr. Erdogan Kaya (Co-PI), Dr. Lori Bruner (Co-PI)

Abstract

The landscape of education is rapidly evolving with the emergence of advanced artificial intelligence (AI) tools such as Gemini by Google, Claude by Anthropic, and ChatGPT by OpenAI. These advancements in Natural Language Processing (NLP) present both opportunities and potential risks for K-12 education, particularly in writing. In response to these challenges, we developed and tested a web-based platform named *Compose With AI*, designed to guide students in grades 4-8 through the process of evaluating AI-generated content and ethically incorporating this content into their science-focused writing. To assess the effectiveness of the *Compose With AI* platform, we employed think-aloud protocols with students and conducted semi-structured interviews with teachers. Our preliminary findings indicate that *Compose With AI* enhances students' abilities to critically evaluate and ethically integrate AI-generated content into their writing. We discuss the implications of our findings for future research and educational practice, considering how this tool could be adapted and implemented across various subjects and contexts. Additionally, we outline plans for wider implementation and scaling of the *Compose with AI* platform.

Introduction

The rapid advancement and adoption of artificial intelligence (AI) tools, particularly generative Natural Language Processing (NLP) tools such as AI chatbots, have created both opportunities and challenges for students and educators alike. In January 2023, ChatGPT, one of the leading chatbot models developed by OpenAI, reached 100 million active users merely two months after its launch, making it the fastest-growing consumer application in history (Hu, 2023). With the platform generating 1.8 billion visits per month, such generative AI tools have the potential to make a measurable positive impact on science teaching and learning in the elementary and middle grades.

Generative AI tools like ChatGPT can assist in building students' background knowledge and explaining complex concepts, which is crucial for comprehending disciplinary texts such as scientific articles (Best et al., 2005; Colwell & Hutchison, 2020). Furthermore, chatbots can generate text models—otherwise known as mentor texts—that can serve as an example for students completing academic tasks (e.g., reporting findings). These capabilities may offer new avenues for supporting student learning from and engagement with scientific content.

However, the educational potential of LLM tools like ChatGPT is accompanied by significant challenges. Students need to possess well-developed critical evaluation skills to accurately assess the content produced by these AI tools (U.S. Department of Education, 2023). OpenAI, the parent company of ChatGPT, has explicitly warned educators about the risks associated with using such tools in educational settings. These risks include plagiarism, over-reliance on AI tools, the generation of harmful and biased content, issues of equity and access, and concerns about the trustworthiness of AI-generated content (OpenAI, n.d.). The unprecedented growth and widespread use of AI applications in education have created an urgent

need for instructional materials and tools that can guide teachers and students in the safe and ethical use of AI.

To address the challenges posed by AI in education, we developed *Compose With AI*, a web-based platform that guides students in grades 4-8 to evaluate AI-generated content and ethically incorporate this content information into their science-focused writing. In this report, we share our preliminary findings and demonstrate the key features of our tool. The *Compose With AI* platform teaches students how to ethically use AI as a writing partner by: (a) guiding students in prompt generation, (b) guiding students in critically evaluating content produced by generative AI, (c) guiding students on the ethical use of AI-generated content in their own writing. Students can use the platform to produce several different types of writing. *Compose With AI* incorporates embedded instructional videos, self-regulated learning prompts (Sanders et al., 2019), and tools for collecting relevant information to generate digital compositions. Through the incorporation of Universal Design for Learning principles (Meyer et al., 2014), *Compose With AI* is designed to be inclusive of all learners.

Research Questions

Our study addressed three research questions:

1. What approaches and strategies inhibit and enhance students' abilities to use and critically evaluate content generated with AI?
2. What uses of AI inhibit and enhance students' abilities to integrate AI-generated content into their writing in ethical ways?
3. To what extent and in what ways do teachers and students in grades 4-8 find *Compose With AI* to be useful for integrating AI-generated content into their science-focused writing?

Compose With AI

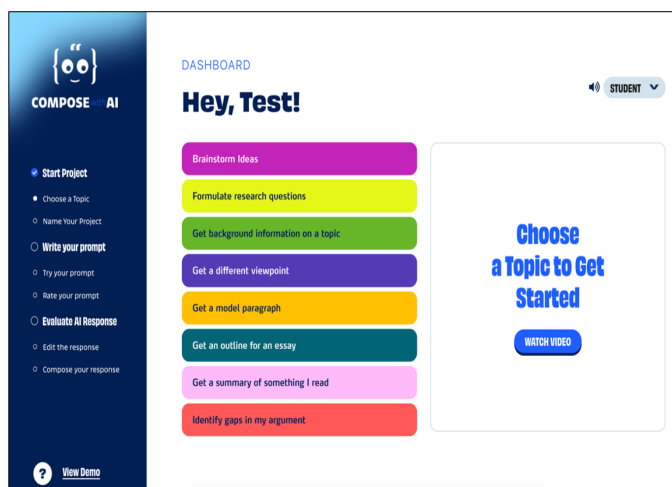


Figure 1. Overview of Student Dashboard

The *Compose With AI* landing page offers two entry points: teacher login and student login. This multi-tiered access structure allows for customized experiences based on user role and needs. As shown in Figure 1, students are greeted with a dashboard presenting a range of topics aligned with science writing (e.g., brainstorming ideas, gathering background information). Each option is accompanied by a brief written description and an instructional video related to the topic. For brevity, we provide a high-level overview of the tool. For a comprehensive exploration, please visit the platform link: <https://www.composewithai.com/>

As an example, when students click the option, “Brainstorm Ideas,” *Compose With AI* employs a structured approach to help students craft effective AI prompts using the R.A.F.T. technique (Role, Audience, Format, Topic). Students are then invited to try their prompts using an AI tool such as ChatGPT and copy and paste their AI-generated responses into *Compose With AI*. The platform then facilitates a critical evaluation process, encouraging students to analyze the content for accuracy, relevance, and ethical considerations (See Figure 2). Students can categorize information as “keep,” “change,” or “omit,” encouraging students to use critical thinking skills essential for navigating AI-generated content.

Figure 2. Critical Evaluation Dashboard

For educators, *Compose With AI* offers a detailed teacher dashboard. This feature allows teachers to monitor student progress, view ongoing projects, and identify areas where additional support may be needed (see Figure 3). Teachers can also review students’ processes and assess the evolution of their prompts. The dashboard presents an overview of class activity, individual student work, and options for providing feedback.

Figure 3. *Compose With AI* Teacher Dashboard

Throughout the platform’s development, we prioritized intuitive visual design and user experience to enhance engagement and ease of use. The color scheme is pleasant and non-distracting, with clear typography and intuitive icons guiding young learners through each step of the writing process. By integrating features that support various stages of the writing process with self-regulated check-ins, *Compose With AI* serves as a tool for enhancing science writing skills while promoting responsible use of AI in science contexts. As we continue to refine and expand the platform based on student and teacher feedback, we believe

Compose With AI has the potential to significantly impact how students approach science focused writing tasks and interact with AI tools in ethical and effective ways.

Methodology

We followed a multi-phase Design-Based Implementation Research (DBIR) process to evaluate the usefulness and usability of the platform and its features. We analyzed data at each phase of the DBIR process as follows: (1) *design phase* (i.e., collected and analyzed teacher interview data about the platform design as the platform was being developed), (2) *delivery phase* (i.e., collected and analyzed student think-aloud data using verbal protocol analysis), (3) *refinement phase* (i.e., refined the platform based on delivery phase data and evaluated our revisions through qualitative analysis of data collected through a second round of student think-alouds), and (4) *retrospective analysis phase* (i.e., after the second round of student think-alouds, all data were combined and analyzed qualitatively to answer the research questions).

During the think-aloud verbal protocol, we presented students with a 45–60-minute think-aloud task that was preceded by a short tutorial on the uses of AI presented in the platform (e.g., brainstorm ideas, get model paragraph) and the evaluation strategies suggested in *Compose With AI* (e.g., lateral reading). During the task, students had access to built-in videos and tips. These videos explain each of the potential uses of AI and evaluation strategies provided within the platform. We utilized a verbal reporting methodology (Ericsson & Simon, 1980, 1984/1993; Fox, Ericsson, & Best, 2011; Pressley & Afflerbach, 1995) to gain access to students' cognitive processes both during the task (concurrently) and after the task (retrospectively). For the task, students: (1) practiced thinking aloud before starting the task to ensure they understood think-aloud expectations, (2) received a prompt (e.g., Choose one animal and explain how this animal's physical features and/or behavior help them survive in their environment?), and (3) followed the steps in *Compose With AI* while thinking aloud about their selections.

Participants included four female students: two students in grade 5, one student in grade 6, one student in grade 7. During the task, we recorded students' screens as they worked and prompted them to think aloud about how they navigated the task using strategies provided by the *Compose With AI* platform. We analyzed these verbal responses to make inferences about the processes participants used, resulting in a verbal protocol. Immediately following the task, we asked students to comment on specific aspects of the task that the researcher wanted to probe further.

In addition to student testing, three teachers reviewed the platform and participated in semi-structured, one-on-one interviews with the research team. The teachers' backgrounds were as follows: (1) a fifth-grade teacher in a self-contained elementary school classroom with 20 years of teaching experience; (2) a seventh- and eighth-grade English teacher with 17 years of teaching experience; and (3) an eighth-grade technology teacher with 21 years of teaching experience. Teachers provided feedback on the platform design and the revisions to the *Compose With AI* platform after the initial round of student think-alouds. During the interviews, we asked teachers three open-ended questions about the following topics: (1) features they thought might work well for students (i.e., the features that might best support students' learning), (2) features that might be difficult for students to use (i.e., features that students might find challenging and/or inhibit their learning), and (3) any features they might add to better support students' learning. We conducted all interviews over Zoom and recorded and transcribed responses.

Findings

What approaches and strategies inhibit and enhance students' abilities to use and critically evaluate content generated with AI?

Results from the student think-alouds show that one of the biggest affordances of *Compose with AI* is that it guides students to proceed in a step-by-step manner by breaking apart the writing process using AI using the following steps: (1) writing an initial prompt, (2) engineering the prompt to better suit the purpose through a series of directed questions, (3) trying the prompt in ChatGPT, (4) evaluating the effectiveness of the prompt by analyzing the relevance of the content generated and revising the prompt if needed, (5) evaluating the credibility and accuracy of the AI-generated content, (6) determining which AI-generated content can be used with modifications, which content needs to be omitted, and which content can be used as it is, (7) composing a response using the information collected throughout the process, and (8) citing sources, including AI. This step-by-step flow was highly valuable for students, with users reporting that it made the process easy and fun. For example, a fifth-grade user stated, “For a class assignment. I mean, it works really well and it’s really fun, I do a nine [on a scale of 1-10]. Well, it depends. If you’ve never done it before and your teacher’s telling you to do it, give it a nine. But if you’ve done it before it’s a lot easier. I’d give it a 10.” None of the participants had previously used generative AI, including ChatGPT, and had not previously been taught how to evaluate online content. The researchers observed that the step-by-step process served as an effective introduction to searching for information and composing a final response. For example, one student reported: “In the beginning, I didn’t quite know what a prompt was, but then I understood,” indicating that the platform itself helps students learn as they go. Furthermore, students who completed a second task provided more critical responses, suggesting the platform becomes easier with repeated use.

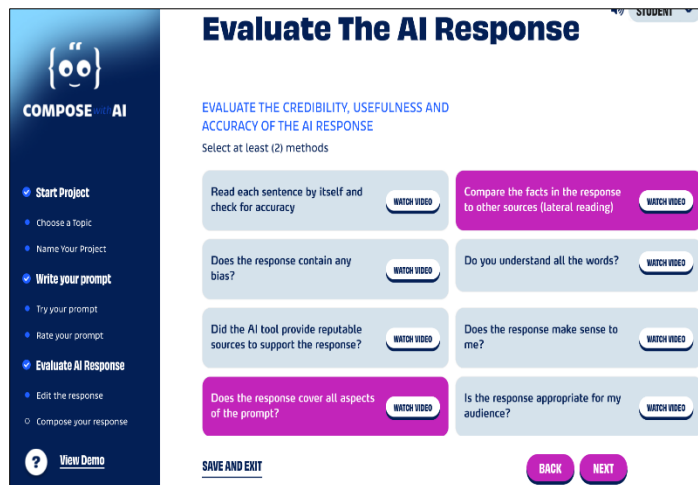


Figure 4. Evaluation Strategy Selections

One of the steps that enhanced students’ abilities to use and critically evaluate AI-generated content was the built-in page that included and explained several strategies for evaluating content and *required* students to choose and use two of the strategies before advancing to the next screen (see Figure 4). After an initial round of think-alouds, we made changes to the screen shown in Figure 4 because students tried to use all of the featured methods, though not all methods are relevant to every prompt. We changed the screen so that students need to select only two methods before moving on. Even the youngest students

with whom we tested the platform found the evaluation strategies to be useful and made comments such as “Some facts are different compared to the other sources” (fifth grade user).

An additional step that was highly useful to students was the page that required students to determine whether they could keep, change, or omit information that was generated by ChatGPT (see Figure 5). In this step, AI-generated content is shown on each screen as the student progresses through the platform. In the user testing, students displayed strong reasoning skills at this step of the process. For example, one student whose task was to use AI to

brainstorm careers related to space chose to omit information about space architects, reasoning that, “... an engineer and an architect are basically the same thing, sort of. Architects make buildings, like they design buildings, and a spaceship is not a building” (fifth grade user; Shown in Figure 5). This step in the overall process required students to carefully read and make decisions about how to use the AI-generated content, which led to effective compositions at the end of the process.

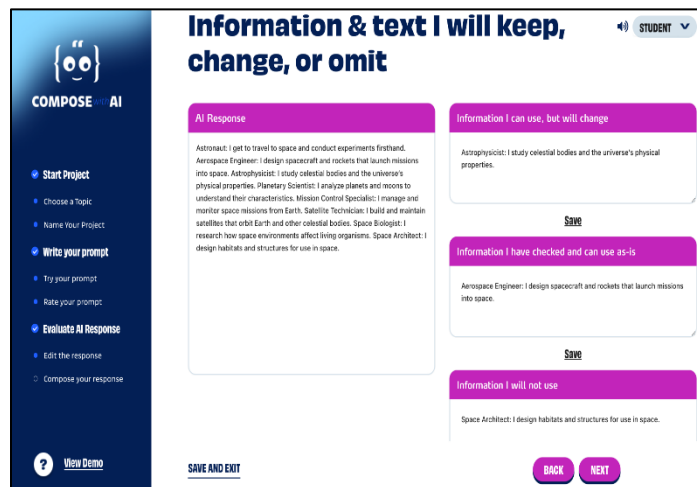


Figure 5. Usage of AI-Generated Content

The primary aspect that inhibited students' progress was that all steps are not self-explanatory. The platform was designed to be used with classroom lessons. Thus, the test users had to be guided through several steps of the process. However, we plan to develop lessons to accompany, eliminating this challenge in the future. Additionally, researchers noted that users did not readily choose to watch the built-in videos explaining the different ways to use AI, nor utilize the built-in evaluation strategies. Thus, future revisions to the platform may include ways to make the videos and evaluation strategies more visible for users.

What uses of AI inhibit and enhance students' abilities to integrate AI-generated content into their writing in ethical ways?

Test users reported that the decision-making screen (shown in Figure 6) helped them decide what to report in their final composition. Researchers also observed that this was a critical step in the overall writing process. Similarly, researchers observed that having user-generated content follow the user throughout the platform was helpful in composing a final response. As Figure 6 shows, when users get to the screen where they compose their final response, the information they have already decided to keep, change, or omit is shown on the left side of the screen. They have the information from the previous step available as they compose their final response. User testing indicated that this scaffolding was helpful because it appeared to reduce cognitive load by allowing the user to complete the task in small chunks rather than composing the response all at once.

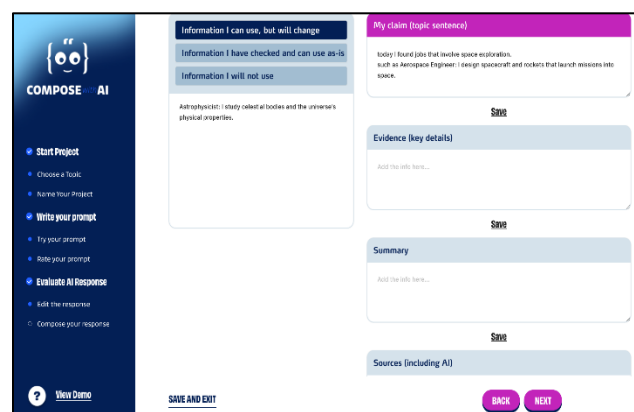


Figure 6. Guiding Students to Compose Their Final Response

To what extent and in what ways do teachers and students in grades 4-8 find *Compose With AI* to be useful for integrating AI-generated content into their science-focused writing?

Three overarching themes about *Compose With AI* emerged through our semi-structured interviews with teachers: (1) advantages of integrating AI-content into instruction in the middle grades, (2) the usefulness and usability of the platform itself, and (3) the affordances of AI-generated content to support students' science writing.

Advantages of Integrating AI-Content with Classroom Instruction

All three teachers noted the benefits of teaching students to use AI tools ethically as part of the curriculum. Every teacher—including the fifth-grade teacher we interviewed—noted that students are entering their classrooms familiar with AI tools like ChatGPT, and that most students have experience using these tools outside of school. As one teacher remarked: “I haven’t been able to find something like this [*Compose With AI*] that helps kids use AI to their advantage; most kids are just finding things online at home that are like, ‘here’s A.I, have fun.’”

All three teachers also noted that they have participated in professional development related to AI tools within the past calendar year, including prior to the start of the 2024-2025 academic year. Teachers reported that these professional development experiences focused on explaining the purpose and function of AI tools, but did not specifically address how to use them to support instruction. As one teacher explained, “This [*Compose With AI*] would be huge for a district that needs a curriculum or a program to teach kids how to use AI effectively.” Notably, none of the teachers we interviewed worked in districts that banned AI usage in school.

Usability of the Compose With AI Platform

All three teachers noted that the *Compose With AI* platform was easy to navigate due to its organized columns with clear headings. One teacher commented: “The tool is kid friendly. It keeps information organized into bite-sized chunks that won’t be too overwhelming for students.” This same teacher further remarked that the graphics and colors were “eye-catching” and would be appealing to students in the middle grades. She explained, “it [the colors and graphics] made me excited to use it.”

Teachers also remarked on specific aspects of the design. While students who participated in our think-aloud protocol did not readily watch the videos embedded in the tool, two teachers took note of them. One teacher noted that she loved the video examples and how we used AI to create the dialogue. She commented: “It felt very fitting for this project.” Another teacher remarked that the videos were short but clear, and she imagined students could “get the gist of what to do without having to sit there for a long time and getting bored.”

A third teacher commented specifically about the “prompting” aspect of the tool. She admitted she had never used AI prompts before and found the step-by-step procedure to be extremely supportive of the process. She noted: “The way it was broken down helped me better think about the types of things you need to say to get the response you’re looking for. This was my favorite part of the tool.”

Using AI as a Writing Tool

The final theme that emerged from our semi-structured interviews was the affordances of using AI-generated content to support students' writing. The teachers noted that *Compose With*

AI provides guidance for many different forms of science writing that are common in the middle grades: narrative and/or informational essays, opinion writing, and inquiry/research reports. One teacher noted that opinion writing is especially difficult for students, and found the “get a different perspective” feature to be particularly useful. She noted that students are “50/50” on their ability to do this, which often hinders their ability to effectively write opinion pieces.

All teachers remarked that *Compose With AI* teaches students that AI-generated tools still require them to have writing skills and craft to use these tools effectively. As one teacher explained, “It [*Compose With AI*] opens my eyes to the idea that AI is a great tool and this [the *Compose With AI* platform] is a good way to show students how to apply critical thinking skills to their writing. It can ‘better’ what they come up with, but they still have to use their own analysis to figure that out.”

Future Directions

The promising results from our preliminary study of *Compose With AI* have opened up several avenues for future development and expansion. Our next primary step is to apply for a new NSF award, such as the DCL: EducateAI initiative under the CSforALL program. This funding would support the expansion of our platform to include the development of a new curriculum. The proposed curriculum will provide structured lessons and activities, ensuring that students and teachers can maximize the benefits of the tool within science classrooms. By creating this robust curriculum, we aim to address the need for guided implementation, as our findings suggest that *Compose With AI* is most effective when used under teacher guidance. This curriculum development is crucial for enhancing the platform’s impact and facilitating its effective integration into classroom settings.

In tandem with curriculum development, we are exploring the integration of more advanced AI features into the platform. One exciting possibility is the inclusion of specific AI agents that can enhance the learning experience. For instance, students could engage in a conversation with a virtual Einstein about quantum theory or participate in an interactive game about acids and bases. These AI-driven interactions could significantly boost student engagement and provide more diverse learning experiences. Relatedly, to improve accessibility and multimodality of the platform, future iterations of the platform may incorporate enhanced multimodality features. By integrating voice recognition and speech synthesis capabilities, we can allow students to interact with AI using spoken language, expanding the ways in which they can engage with the platform.

While our current focus is on science-focused writing, we recognize the potential for *Compose With AI* to support learning across various disciplines. We plan to explore adaptations of the platform for subjects such as computing, engineering, and social studies, customizing the core features to support coding, writing, and critical thinking in different academic contexts. This cross-disciplinary approach could significantly broaden the platform’s impact on student learning.

Another area of potential development is the integration of AI-powered tutoring features. By incorporating feedback mechanisms that provide students with personalized suggestions based on their prompts and writing, we can create a more responsive learning environment that adapts to individual student needs such as emergent multilingual learners.

As we pursue these future directions, we remain committed to prioritizing ethical considerations, user experience, and alignment with educational best practices such as UDL principles in all our development efforts. The potential of *Compose With AI* extends beyond its current form as a tool for science-focused writing; we envision it evolving into a more interdisciplinary platform for cultivating the AI literacy skills that will be essential for students’ future success. By continually refining and expanding the *Compose With AI* platform based on research findings and user feedback, we aim to create a versatile and effective tool that supports students in their immediate academic tasks and augments their writing skills and prepares them for a future where AI literacy will be a fundamental skill.

References

- Best, R. M., Rowe, M., Ozuru, Y., & McNamara, D. S. (2005). Deep-Level Comprehension of Science Texts: The Role of the Reader and the Text. *Topics in Language Disorders*, 25(1), 65-83. <https://doi.org/10.1097/00011363-200501000-00007>
- Colwell, A., Hutchison, A. & Woodward, L. (2020) *Digitally-Supported Disciplinary Literacy for Diverse K-5 Classrooms*: Teachers College Press.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87(3), 215-251. doi:10.1037/0033-295X.87.3.215
- Ericsson, K. A., & Simon, H. A. (1984/1993). *Protocol Analysis: Verbal Reports as Data*. MIT Press.
- Fox, M. C., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting Methods. *Psychological Bulletin*, 137(2), 316-344. doi:10.1037/a0021663
- Hu, K. (2023, February 2). ChatGPT sets record for fastest-growing use base. Reuters. <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>
- Meyer, A., Rose, D. H., & Gordon, D. (2014). *Universal Design for Learning: Theory and practice*. CAST Publishing.
- OpenAI Platform. (n.d.). *Educator considerations for ChatGPT*. <https://platform.openai.com/docs/chatgpt-education>
- Pressley, M., & Afflerbach, P. (1995). *Verbal Protocols of Reading: The Nature of Constructively Responsive Reading*. Lawrence Erlbaum Associates.
- Sanders, S., Losinski, M., Parks Ennis, R., White, W., Teagarden, J., & Lane, J. (2019). A Meta-Analysis of Self-Regulated Strategy Development Reading Interventions to Improve the Reading Comprehension of Students with Disabilities. *Reading & Writing Quarterly*, 35(4), 339-353. <https://doi.org/10.1080/10573569.2018.1545616>
- U.S. Department of Education, Office of Educational Technology, Artificial Intelligence and Future of Teaching and Learning: Insights and Recommendations, Washington, DC, 2023. <https://www2.ed.gov/documents/ai-report/ai-report.pdf>