From Ban to Beta: Reimagining AI's Role in Program Evaluation Pedagogy

David T. Marshall
Auburn University

When news broke of ChatGPT's release in the fall of 2022, I was skeptical of the new technology. Recognizing that generative artificial intelligence (AI) enhances its performance with increased usage by individuals, I resolved to abstain from contributing to its development. My concerns ranged from the unknowns associated with computers having this much capability and knowing it would only 'improve' to concerns about protecting data. About 5 months later, I changed my mind completely. As a faculty member, I teach statistics, research methods, and program evaluation in a college of education. ATLAS.ti, a qualitative research software, released its AI Coding beta in late March 2023 (ATLAS.ti, 2024). This was a significant moment in terms of how I viewed AI. I quickly reasoned that since I teach graduate students how to conduct research and analyze data, it was going to be important for me to understand how to use these tools and teach students to use them in ways that: (a) are ethical and (b) that enhance—not replace—their research capabilities.

Description of Course

The first course in which I integrated AI was my Program Evaluation course. The students are generally second and third-year doctoral students from a range of programs, including adult education, counseling, educational leadership, educational psychology, kinesiology, nutrition, and special education. The semester-long course meets once per week for 3 hours. Throughout the semester, students work towards the creation of an evaluation plan for a real client. The final deliverables include a roundtable presentation to the class and a final paper outlining the evaluation plan they created. Most of the assignments throughout the semester have students working towards the development of this plan. Each class contains some lecture, but most of the time is devoted to discussion and applied activities.

How the Treatment of AI Has Changed

Beginning with the summer 2023 semester, I added an AI clause in all course syllabi, including for my Program Evaluation course, that explicitly banned the use of AI without specific instructor permission. It's worth noting that I never permitted students to use it. Although I did start discussing the use of AI in research, for example, highlighting the AI Coding feature in ATLAS.ti. I reasoned this was appropriate; my students should be aware of the range of research tools that exist. However, since

they were only creating an evaluation proposal, and not collecting and subsequently analyzing data in my class, I did not need to allow them to use AI for their evaluation proposals. Several students also take a practicum course with me following the semester where they actually implement their proposal, but that was not the requirement for the Program Evaluation course. And, as such, there was no data to be analyzed.

This was my policy until the fall 2024 semester, when I decided to take a different approach towards AI. I no longer ban the use of AI tools, but I did begin requiring an "AI Statement" to be included in every assignment. The statement needed to disclose whether AI tools were used in the process of completing the assignment, and if so, what platform was used (e.g., ChatGPT, Microsoft Copilot, Grok, Google Gemini), what prompt was used, their aims for using the tool(s), and whether it helped them to achieve their desired outcome. I decided that we were going to still learn about how to evaluate programs and policy in the course, but we were also going to take the opportunity to learn how to use emerging AI tools together.

I would also include AI guidance as part of the assignments, suggesting when AI might or might not be helpful. For example, AI would not be particularly useful for an assignment that asked students to reflect on an activity, since that is personal.[1] Students were not required to use AI, but they were required to disclose when they did. Students who elected not to use AI were still required to include a sentence stating as such. Anyone who submitted an assignment without an AI statement would have it returned ungraded until they resubmitted the assignment with the statement included.

When I reviewed assignments I had graded with the class, I would begin with a review of how students in the class used AI to complete the assignment, and whether it was helpful. For example, the first assignment of the semester asked students to work in groups to create a presentation about an assigned evaluation approach. Fewer than half of the students stated that they used AI. One student later shared with me that this was perhaps because they were skeptical that I was serious when I said it was alright for them to do so. I was among the first in our college to approach AI this way, and students who were thought to have used these emerging tools in other classes were penalized for doing so.

Students who used AI for their assignments did so in creative and productive ways. A couple of groups used ChatGPT to come up with a scenario they used for a group activity. Their prompts must have been similar because both groups produced similar activities that involved a public health-related example. Another group shared that they used Copilot to create a simple description of the evaluation approach on which they were presenting. And one enterprising student decided to use ChatGPT to create a rap, complete with music and a beat, that was about the CIPP model (Stufflebeam & Coryn, 2014), her group's assigned evaluation approach. It was as ridiculous as it was catchy, and was

1 I should credit Dr. Asim Ali, director of the Biggio Center for the Enhancement of Teaching and Learning at Auburn University, for this suggestion. I would not have thought to do this if not for a conversation we had about AI and teaching in higher education.

something that would have previously been unheard of.

Moving Forward

The pivot in my approach to AI was welcomed by most students. However, when I reflect on the fall 2024 semester, I realize that we only scratched the surface regarding how these tools could be used. I permitted students to use AI, but I never offered guidance on how they should go about it. The extent to which one is successful using AI is often mediated by their prompt engineering skills (Alammar & Grootendorst, 2024). Prompt engineering refers to one's ability to draft prompts for generative AI tools. Being able to craft clear and specific prompts with an appropriate amount of context for the task at hand usually leads to more useful and sophisticated AI responses. Similarly, knowing when to use which AI tool, or when to use a given model within an AI tool (e.g., understanding when ChatGPT-4o or GPT-o3 is more useful, or when using the Deep Research tool would be beneficial). Different large language model platforms will have differing strengths, and understanding when one is more useful for a given task is important.

This fall, 2025, I plan to spend more time having students use a range of generative AI tools and collectively learning to get better at prompt engineering. For example, I plan to have students work in groups to develop a prompt to respond to a scenario. We will then share the prompts and the results, learning tips for crafting better prompts along the way. Once we decide which prompt was the best, I will assign each group to use a different AI tool using the model prompt. For example, one group might be assigned to use ChatGPT, another to use Google Gemini, another to use Grok, another to use Perplexity, and another to use Microsoft Copilot. We will then debrief again, with each group sharing their screen to show what they uncovered. Another version of this might involve having each group try three of the five suggested models to compare results.

Concluding Thoughts

The shift in my AI policy was meant to allow us to learn how to use these tools collaboratively, and I believe we made meaningful progress in that direction. While I have more expertise in program evaluation than my students, we are on equal footing when it comes to exploring the potential of AI tools. Generative AI is likely here to stay, and as such, we must equip students with the ability to use it productively. As Ethan Mollick (2024) reiterates in *Co-Intelligence*, the only way to understand AI is to use it.

Acknowledgements

I would like to thank Katelyn Nelson (Auburn University) for the thoughtful feedback and comments she provided on previous versions of this draft. It is a better product as a result of her insights. Any flaws that remain are mine alone.

References

Alammar, J., & Grootendorst, M. (2024). *Hands-on large language models: Language understanding and generation*. O'Reilly Media.

ATLAS.ti. (2024). *Introducing: AI Coding Beta powered by OpenAI*. Author. https://atlasti.com/ai-coding-powered-by-openai

Mollick, E. (2024). *Co-intelligence: Living and working with AI*. Portfolio.

Stufflebeam, D. L., & Coryn, C. L. S. (2014). *Evaluation theory, models, & applications* (2nd ed.). Jossey-Bass.