

A Personal Reflection on Teaching Ethics with AI: Learning Through Experimentation

Roger Martin
University of Virginia

If you had told me a year ago that one of the most energizing parts of my graduate ethics course would involve artificial intelligence ("AI"), I would have been quite skeptical. I teach in a one-year master's program for accounting students – bright, serious students who appreciate clarity, structure, and objectivity. So, when I introduced AI into a small, discussion-based class on business and accounting ethics, I wasn't entirely sure what to expect. What I found surprised me, challenged me, and ultimately changed the way I think about teaching ethics and teaching more broadly in an AI world.

I'll admit, I was nervous. I didn't know how the students would take to AI, nor was I confident that I understood how AI might fit in this type of theoretical, very sensitive course. I tried to model my own uncertainty about AI throughout the course. I shared moments when I wasn't sure what to expect from AI, or when a response surprised me. I let the students see my process of trial and error. Instead of presenting AI as either a threat or a solution, I framed it as a tool we were all learning to work with. This modeling of uncertainty seemed to resonate with students.

And to their credit, they were willing to experiment with me. They ended up diving into the activities enthusiastically and were eager to share their discoveries with each other and with me. After the course ended, I got messages from some of them telling me how they are using AI in their jobs, and how they feel more capable and comfortable than many of their peers because of the practice they got during the course.

I feel lucky to have enjoyed this success, and I'm glad to share my experience with others to encourage their own experimentation.

Activities to Explore AI

During the semester, I implemented three core AI activities to help students engage with AI from an ethical perspective. I wanted those activities to develop students' thinking and their ability

to assess how helpful AI could be in ethics and ethical decision-making contexts. In hindsight, I think the progression of activities – moving from using AI as a tool to learn ethics, to making AI the focal point of our newfound ethical lens – was an effective (though admittedly lucky) design choice for the course.

Comparing AI Responses to Our Own

Our first AI-integrated activity involved a case we had previously worked through in class. After the discussion, I shared sample responses generated by an AI tool when I asked it to apply each of the ethical frameworks we had covered up to that point to address the case discussion prompts we had used in class.[1] The answers were polished, articulate, and filled with ethical terminology we had been using, so at first glance, the students were impressed. But then we started asking harder questions. Was the actual ethical recommendation clearly stated? Did the response include a logical explanation of the ethical reasoning? Was that reasoning internally consistent? Almost always, the answer was no. It helped students realize that impressive vocabulary doesn't equal sound reasoning. They began to see how easily they might be swayed by the appearance of expertise, even when the substance was thin.

This exercise also provided a safe way to critique responses without the personal risk that comes with evaluating each other's work. By assessing AI-generated answers, students were indirectly learning to be more critical of their own thinking. It also helped me communicate what I value in student responses—clarity, specificity, and consistent application of ethical frameworks. And because the AI outputs were so clearly written, they also demonstrated the value of using direct, active voice in communicating ethical reasoning.

Using AI to Explore and Develop Our Understanding

In the second AI activity, I asked students to prompt the AI tool to take on the persona of an expert in one of the ethical frameworks we had studied and engage in conversation with that "expert" about a case they had prepared for class, but that we had not yet thoroughly discussed. Students found this activity playful and revealing. They were able to push the AI, pose alternative case scenarios, and press the AI tool to explain how it balanced potentially competing perspectives, which is something we always find challenging in our own discussions. Most students found the AI tool could explain concepts in a way that was clearer and more detailed than our textbook, making the interaction not just engaging but also educationally valuable. Students discovered a fun extension of this activity by changing some of the fact patterns within the case to see how it affected the expert's evaluation; some of those revisions became comical, but students were learning to test the boundaries

1 The cases used in these AI activities were written by me, and I made it clear to students that they had my permission to use them in this way. We discussed the potential inappropriateness of providing copyrighted material to AI tools, so I was careful to explicitly grant permission to use my own written materials for this purpose.

of some of the important positions within these frameworks, in a way that was fun and instructional.

One thing I didn't anticipate was how much these AI interactions would humanize the subject matter. Ethics can feel abstract and theoretical, especially for students grounded in technical disciplines. But when students watched the AI stumble, contradict itself, or hedge, it seemed to give them the confidence to do the same in our subsequent discussions, without feeling like they had to have the "perfect" answer. I thought class discussions became more authentic and inclusive after this activity, and students were more open and experimental in their contributions. It gave them space to be learners rather than performers.

Making AI the Subject of Ethical Scrutiny

Later in the semester, I assigned student pairs to explore ethical challenges posed by AI itself – topics like bias, misinformation, labor exploitation, and environmental impacts. I hoped the project would encourage students to think critically about the real-world implications of AI technologies in ethical contexts. I asked each pair of students to brief the class on their topic and then lead a discussion about how the ethical frameworks we had studied could be applied to evaluate and mitigate these challenges. What stood out to me was the gravity of the issues that students faced, particularly as they struggled with balancing the ethical dimensions of the problems. But I think it turned out that the struggle was itself the lesson. The ethical terrain of AI isn't simple, and pretending otherwise would be a disservice.

It also prompted deeper discussions than I had expected. In one case, a student group talking about misinformation and deepfakes found themselves genuinely unsure how to balance free expression against potential harm. Another group examining environmental costs of large language models wrestled with how to weigh long-term sustainability against short-term utility. These weren't academic debates; they were grounded in real tensions the students felt.

We also talked about the ways AI might shape the expectations of future professionals. If clients or employers come to expect instant, polished analyses, will there still be space for deliberation, uncertainty, and doubt? These conversations reinforced for me the importance of preparing students to recognize the value of, and preserving the space for, human judgment.

The Experiment Worked Better Than I Could Have Hoped

Looking back, I'm struck by how willing my students were to jump into something unfamiliar – and how much they taught me in the process. I learned that AI can be a mirror, reflecting the strengths and weaknesses of our own reasoning. I learned that it's okay to invite students into an evolving space, to model uncertainty rather than mastery. And I learned that experimenting with new tools isn't a distraction from learning – it's fun and it clearly improved learning in my class.

This experiment is only the beginning of AI use in my classroom. I plan to further refine how I integrate AI into my teaching and to be more intentional about guiding students in how they can use AI to support and improve their learning. I want to help students see AI as a supportive tool – one that can make their studying more efficient and their analysis more impactful, but *not* one that replaces their need for critical thinking and deep engagement with the material.

And from a purely personal perspective, this experience has opened my own eyes about AI. Not because AI has all the answers, but because it helps us ask better questions. And sometimes, that's the best place to start.