Emphasizing Process Over Product in Student Writing

James McGrath

Butler University

There are a lot of tactics that I have adopted to discourage inappropriate use of AI, and to make it unlikely that a student will do well if they nevertheless choose to go that route. One is having students write their work in a shared document that I set up. I then have ready access to the process and not merely the product, since cloud storage services save backups that are timestamped. This makes it easy to see if a student goes from having written no words to a full-length essay in the hour before a deadline, but that is not the greatest advantage. It places the focus on the process, allowing me the opportunity to provide input along the way at times. Even when that is not feasible, the message is still sent that it is the path of investigation and discovery that is most important.

Students assume that what we are looking for is something polished and perfect that grants them an "A." We grade what students submit, sending the message that what matters is the end result and not how they got there. To adapt the saying "show me your budget and I'll show you your values" for this academic context, "show me what you assign a grade to and I'll show you what is important in your class." Of course, many academics would insist that there are things that we care about that aren't graded, things that perhaps cannot be graded. The question that we rarely ask ourselves is why then we assign points to the things we do. In my own case, the reason I assign a novel to be read is the experience itself of personal transformation that happens when reading it. Grading personal transformation and engagement is difficult, and so once upon a time I might have resorted to quizzes on details from the reading, just to try to make sure the reading is done. I stopped doing that because it conveys that the point of doing the reading is to remember the names and plot details.

The same thing can be accomplished by asking for a personal reflection that requires students to talk about specific assigned pages of reading. This is more meaningful and also addresses the issue of reliance on AI (not to mention Sparknotes and other issues that existed before the release of ChatGPT). An LLM-powered chatbot is unlikely to ever be able to deal with specific pages from a specific edition of a copyrighted text. An LLM imitates human speech in very plausible ways. It has no mechanism for identifying and vetting information. Its training will allow it to generate plausi-

ble output that fits a discussion of a specific novel or textbook, but not in a way that identifies precise pages or can answer non-leading questions about what an author wrote or what characters say and do. Even before AI, a student could make things up and hope they get away with it. Now in the era of AI, they are more likely to get a chatbot to generate a plausible-sounding and grammatically perfect answer. Either way, the shortcomings and inaccuracies will become apparent. You can grade what is submitted without needing to determine whether the fabrication was done by the student themselves or outsourced to someone or something else.

Even more useful and generally applicable are things that I now increasingly do at the front end of the process. Students do not typically understand what we are looking for from their work, nor what the point is of assigning it. Explaining to students the point of what I am requiring them to do and showing them why reliance on a chatbot will not serve them well is far more effective in my experience than seeking to catch them later. Moreover, in an ideal world I would not want to catch them; I want to persuade them to make better choices. One way of doing that is to take the pressure of the grade off, so that it is not the final product alone that determines how they do in the class. If students must get an A to keep their scholarship and remain a student, even if they are inclined to be honest, they may feel it is too risky not to turn to AI to do the work for them.

The unexamined premise is that AI *can* do a better job than they can. Its grammar may be more polished, and that too is a pressure valve we can easily release. The most important step, however, is that they see for themselves how an LLM produces unsatisfactory output. I realized this during the semester when ChatGPT was released. I decided to tell students to use it, explore what it can and cannot do, so that together we learn about this new technology. They could earn a small number of points for just asking it things and submitting the output as an illustration, and much more by engaging with it in a more meaningful interactive way. One student asked it whether Mary Magdalene was a prostitute. The chatbot gave a good answer, that this is a widely-held view not supported by our early sources about her. The student asked for sources and it offered some. She asked for links and it gave URLs. None of them worked and she pointed this out. The chatbot apologized, explaining that things sometimes move on the internet. The student then began looking for the sources and discovered that while author names were for the most part real (one of them had been a guest speaker in the class that semester!) The titles of books were not things these authors had written. As a result, the student understood that an LLM generates plausible text but has no mechanism to ensure that the output corresponds to things that exist in the real world. The process of learning this for themselves conveyed it in a way that a professorial warning would not have accomplished nearly as effectively. Taking students through this process in a live classroom demonstration can likewise be effective. Whether AI output is likely to have accurate information has nothing to do with how easy or difficult the question is, but how clearly the information is woven into relevant patterns of speech in the training data. Most students (and alas, some academics) still do not grasp this.

We who write and publish know that the only way to produce strong content is to finish a draft well ahead of the deadline, set it aside, then reread it later and make revisions. Students do not know this. Even when first-year seminar courses require that students write a draft and then revise it, they may not understand *why* it is required, nor how to best accomplish it. I have taken to talking them through the steps of setting a calendar reminder long enough before a deadline to ensure that there is time to write, set aside, reread, and revise. They do not typically use their calendar in this way; thus they are caught off guard by approaching deadlines too late for them to produce quality work even if their research and writing skills are good.

Having AI revise one's own writing is sometimes permissible, but no automated process will make only revisions that are guaranteed to be improvements. A grammar tool will suggest things that should be rejected based on our human comprehension of the language. Asking a chatbot to take a draft and "improve" it suggests the user thinks AI understands what they have written and knows what constitutes an improvement. Neither is true. To illustrate this, I asked an LLM to "improve" a sample of writing taken from this or that popular young adult novel that students will be familiar with. It is easy to show the changes the LLM makes by pasting the original text and the version revised by the LLM into separate Word documents and then using the compare documents tool in Word. The changes are mechanical, unnecessary, and will include changes to the meaning as well as an occasional improvement. I show this to students, and they get it.

I also show them what it looks like when I compare two versions of a document that I have written. The pattern and character of changes made by a person vs. an LLM look almost nothing alike. Students grasp what I look for in their revisions, how an automated process falls short, and that I can also use tools like this to catch academic dishonesty. It is satisfying as an educator that this dissuades misuse of AI more often than not, allows me to catch inappropriate usage of AI without policing it per se, emphasizes the learning process rather than the output, and helps students learn not just about course content but about a technology that is surrounded by hype and pitfalls.