



## ARTICLE

# Using ChatGPT to Teach the CREAC Format to First-Semester Legal Writing Students

**Sarah Parks**

*Visiting Clinical Assistant Professor of Law  
Indiana University Robert H. McKinney School of Law*

Published: December 2023

I am both a legal writing professor and a language student—I am learning to speak German. In German, some nouns are feminine, some are masculine, and some are neuter. Why? For seemingly no reason at all. This non-explanation is hard for me to accept. First-year law students, too, are learning a new language. In the same way I felt frustration with German gendering, my students felt skepticism, frustration, and doubt in the face of the new norms and expectations I asked them to follow in our first-year legal writing course. For example, they resisted the CREAC format,<sup>1</sup> rebuffing it as exactly the type of archaic nonsense they were warned they'd find in law school.

Unlike the gendering of German words, however, there are good reasons to follow CREAC. I brainstormed a couple of ways to illustrate its utility. I could, of course, write examples of poor formatting for the students to critique, but I feared the students would view my sample as tainted by my pro-CREAC bias. Ultimately, I turned to a new tool at my disposal, which would generate samples for me in an objective (and still probably deficient) manner.

---

<sup>1</sup> CREAC stands for Conclusion, Rule, Explanation, Application, Conclusion.

With cheer and optimism, I set out to prompt ChatGPT<sup>2</sup> to generate samples for my first-year students to critique; this assignment would literally write itself! As you might expect, the preparation instead took quite a bit of work. In the end, however, the experiment was a success. ChatGPT provided some truly abysmal examples of memo writing, ripe for even 1L critique.

My hope was that in approaching the familiar facts and law from a new perspective (that of a reader), the students would begin to value the CREAC format. I also hoped that the exercise would deter students from using ChatGPT in their classwork. Ultimately, I believe this activity achieved both ends. The students were highly engaged and interested in reading the ChatGPT samples. As a group, they concluded that ChatGPT was not a reliable tool in this context. They also expressed a greater understanding and appreciation of the CREAC format.

## 1. Preparing the Assignment

To develop AI-generated samples, I provided ChatGPT with a short fact pattern and one case excerpt. Both related to a claim for Negligent Infliction of Emotional Distress (“NIED”). The students had been working with these same materials for two weeks. I prompted ChatGPT to analyze whether the client’s perception of an accident involving a hydraulic press was sufficiently first-hand to state a claim under the law I had provided.

Right off the bat, ChatGPT identified the wrong law. ChatGPT should have analyzed the facts using a foreseeability test. Instead, ChatGPT focused on a “zone-of-danger test.” The caselaw I had provided expressly rejected the zone-of-danger test. To correct this error, I supplied ChatGPT with a second case excerpt, which restated the correct rule.

ChatGPT provided a new, but still inadequate, answer. It did not abandon its discussion of the zone-of-danger test, but it did *add* the correct foreseeability test. Curiously, it also added a third test, the “physical impact test,” which did not appear in either of the cases I had provided. Nevertheless, I was pleased with this outcome. Like ChatGPT, my students had some difficulty picking out the correct test, and this would be good for them to revisit.

On top of its inability to identify the correct law, ChatGPT had also hallucinated facts. It described the client as having witnessed an airplane accident. I have no idea where this came from. I had not provided ChatGPT with any

---

<sup>2</sup> I used the free, public version of ChatGPT. Other generative artificial-intelligence tools, particularly those geared towards legal work, would likely yield more practice-appropriate results.

materials relating to an airplane accident. To remedy the fact hallucination, I again provided ChatGPT with the fact scenario concerning the hydraulic press and asked it to update its analysis. Its revised analysis was reasonable, albeit threadbare, and the facts it referenced were correct.

I asked ChatGPT to organize the analysis in the CREAC form, and it did so. I reran the answer a few different times to produce three different samples.

## 2. In-Class Work

My students had been studying the same NIED fact pattern and case excerpts for a few weeks. In a prior class session, I had asked the students to write their own CREAC analysis of the issue, and I had also provided them with a good, human-generated sample for review. They were, therefore, familiar with the analysis I had asked ChatGPT to generate.

In class, I divided the students into three groups. Each group read all three samples and was assigned one sample to analyze and present on. I asked each group to consider the following prompts as they critiqued their sample:

- Are there issues with the CREAC form?
- If there were headings, were they descriptive?
- Was there a conclusion? Was the conclusion correct?
- Was the rule correct?
- Was the explanation section adequate?
- Did the application section draw comparisons to the precedent cases?
- Was the application section what it should have been?
- Were any facts included that should not have been? Were any omitted that should have been included?
- Did the sample consider any weaknesses or counterpoints?
- Did the sample restate the conclusion, and did the conclusion match the initial conclusion?
- Was the tone appropriate?

Through the group presentations, the class identified many issues with the samples. The main three issues, paraphrased, were (1) ChatGPT's failure to reach any conclusion, (2) ChatGPT's lack of distinct Explanation and Application sections, and (3) ChatGPT's use of imprecise and space-filler language.

### *2.1. Students experienced frustration as readers because ChatGPT did not lead with a conclusion (or indeed, reach one at all).*

The students quickly identified that, despite having a heading entitled “conclusion,” ChatGPT had failed to reach a conclusion in any of the samples. It can be tempting for first-year law students to do the same. Due to the test-taking strategies they learn for essay exams, students tend to “info dump” rather than provide precise analysis. They are tempted to leave all options on the table rather than tying themselves to just one conclusion. While this might be a successful test strategy, it is an ineffective way to write.

ChatGPT’s work product illustrated that legal writing devoid of any conclusion is difficult to follow. The students experienced discomfort and disorientation in reading these samples. We explored how stating the conclusion first anchors the entire analysis. The lack of conclusion had a waterfall effect: the Explanation and Application sections both felt random and arbitrary because they were flowing neither from nor to any centralized point. Approaching the piece as a reader allowed the students to see the value of a clear conclusion up front, which had been hard for them to appreciate in the abstract.

## *2.2. Students better appreciated the different purposes of the “E” and “A” sections.*

My students had demonstrated confusion distinguishing between the Explanation and Application sections within the CREAC format. They expressed that, in their view, it made better sense to move from the Rule statement straight into the Application. This perspective, however, was that of a writer, not that of a reader. As a writer, they felt confident. They had lived and breathed the fact pattern and caselaw for a few weeks. A reader, however, would not be as educated on the concepts. By experiencing the analysis as readers, the students understood better that the Conclusion, Rule, Explanation, and Application sections all served a distinct, logical purpose.

In the samples it generated, ChatGPT did not thoroughly analyze the caselaw in the Explanation section. Indeed, one sample did not contain *any* explanation whatsoever. Even where ChatGPT did include some discussion of the caselaw, it seemed to cherry-pick facts at random from the cases. The students highlighted the omission of material facts and questioned why other facts were mentioned at all. Consequently, the Application section felt arbitrary. There was no common thread connecting the Rule language with the facts used in the Application section. As I had hoped, the students began to appreciate that unless the Explanation section illustrated the Rule, the Application section could not draw logical analogies.

This concept was well worth the significant amount of class time we devoted to it. As I had hoped, the students left with a much greater sense of buy-in to the CREAC form due to the disorientation they felt navigating ChatGPT's Explanation and Application sections.

### *2.3. ChatGPT mimicked lawyerly language, detracting from the substantive analysis.*

One class discussion point I had not planned on was about ChatGPT's tone and phrasing. Partly to avoid taking a position, ChatGPT's wording was mealy-mouthed and, in some instances, just odd. It wrote with throat-clearing language, nominalizations, and passive voice. The tool was clearly trying to mimic the lawyerly tone but failed to provide any lawyerly reasoning. We discussed as a class the importance of not hiding behind legalese to avoid providing substantive analysis.

We discussed why ChatGPT's manner of writing was unsuitable for legal work. At best, the writing was boring and wordy. At worst, it presented a misleading view of the law. For example, in one rule section, ChatGPT generically stated that the court would "consider these and other elements." Though not the intent (can AI have intent?), this phrasing would signal to the law-trained reader that (a) certain rule language was omitted, and (b) "elements" were in play. Neither was true. The students did not pick up on this issue on their own, but after a discussion they appreciated the problem and its implications for their own writing.

First-year law students often speak imprecisely to fill the space, to sound knowledgeable, or to avoid taking a hard position. ChatGPT, which employs many of the same techniques, cannot alleviate these common problems for students. We discussed how the students should avoid ChatGPT's error of accidentally signaling something other than the intended meaning and should avoid using words that are legal terms of art. Substantively, the danger that ChatGPT and first-year students pose is the same: neither knows what they don't know. Neither can effectively check the other for substantive correctness; students will have to rely on traditional research methods or more advanced tools.

When planning this exercise, I did not expect ChatGPT's omissions or imprecise wording to be part of the discussion. On the contrary, I had expected to run into issues with ChatGPT's notoriously confident tone. If it was wrong, I expected it to be assertively wrong. Nevertheless, the exercise ushered in a conversation about the dangers of imprecise, space-filler language in legal writing.

## 3. Reflection

This activity was worthwhile for multiple reasons, but I had two main takeaways. First, my students got to explore the (non)feasibility of using ChatGPT in their work, at least with the limited knowledge they currently have and ChatGPT's current capabilities. Second, the students got to apply their knowledge of CREAC in a tangible way. They sat as the reviewer over familiar facts and law. Approaching the CREAC format as a reader allowed the students to see the importance of a solid conclusion, precisely worded rules, and distinct, thoughtful "E" and "A" sections.

While I prompted ChatGPT to present its answers in the CREAC format, it would be interesting to try the same exercise without tying ChatGPT to the CREAC format. Given how deficiently it followed the directive, however, allowing ChatGPT free rein might yield responses too far afield to allow for adequate critique in the time I allotted for this exercise.