# Group Exercise:

# Best practices in using AI Language Models for Writing & Editing

**Background**

There has been a lot of discussion surrounding large language models, or “AI”, for writing. While these models have many obvious flaws, several 北大 students told me they already find them useful for editing text. In addition, these tools will likely only keep improving over the next few years

The goal of this exercise is for you to assess whether AI tools like ChatGPT can help you write better scientific abstracts and, if yes, to develop a set of best practices.

*Note: The goal of this class is to prepare you to write scientific articles under realistic conditions. You are allowed to use AI programs on your final assignment, for example to edit text or to generate a first draft which you then modify by hand. However, you need to properly attribute all use of these programs. If you use AI-related tools, you must indicate in the acknowledgements which tools you used and which parts of the text you edited.*

*Using a program like ChatGPT without attribution is plagiarism, which results in an automatic failing grade for the class and might also entail further disciplinary action from the university.*

**Instructions**

First, read the blog post “How to use ChatGPT in scientific writing” by Stephen Heard.

Second, form groups of about 3 students. Each group needs to have at least 1 laptop with access to an AI chat program (e.g., ChatGPT or Google Bard). If you have never used an AI chat program before, take a few minutes to play around and familiarize yourself. You should quickly start noticing that the program can produce wrong or nonsensical information. For example, in 2023 ChatGPT reliably failed this question: “2 students leave Beida together and arrive at Tsinghua after 30 minutes. How long would it take for 4 students?”

Third, here is an abstract from your Chicken midterm papers:

During chicken preparation, flavor enhancement and bacterial safety play an important role. Our study introduces an innovative approach aimed at enhancing flavor and reducing bacterial contamination in cooked chicken and give the optimum combination of heating time and temperature. Traditional chicken preparation methods, including raw consumption, have long been associated with flavor challenges and bacterial safety concerns. Despite innovations like improvements in slaughtering techniques and bird size, these issues persist, necessitating novel solutions. This study addresses the need for a chicken preparation method that simultaneously improves flavor and reduces bacterial contamination, offering a safer and more palatable alternative to traditional approaches. Here we show that our novel chicken preparation method, involving seasoning and heating, significantly enhances flavor and reduces bacterial contamination. Our study reveals that contrary to previous beliefs, chicken flavor can be improved while ensuring bacterial safety through our innovative method. Additionally, our findings contribute to the advancement of cuisine realm by establishing a standardized approach to chicken preparation. The development of our novel chicken preparation method not only addresses longstanding challenges in cuisine but also has broader implications for public health, offering a safer option for poultry consumption. Moreover, our approach underscores the importance of integrating culinary expertise with scientific knowledge to solve complex food-related issues.

Here are some of the issues we noted while grading:

Second sentence – “give” should be “gives” or “provides”

Sentence starting with, “Here we show…” – Significant overlap with sentence 2. Better to restructure the abstract.

Last sentence – How does this work integrate culinary expertise with science? Nowhere else in abstract is culinary expertise mentioned.

Overall – The abstract states the result first, then the motivation. That’s a good structure for a short conference presentation but can backfire for many scientific journals. In addition, the first few sentences are unnecessarily repetitive. Better to have a single sentence that clearly states the research gap. I think this abstract would be better if it followed the classic Nature model from class.

As a group, design a set of experiments to see how an AI program can help you improve this abstract. To do so, first come up with a methodology for how to evaluate and compare two texts (for example, have everybody in your group read the original and modified text, then vote on which one is better; or, ask ChatGPT to evaluate the text for you, then compare whether your group agrees).

Next, discuss and experiment which input prompts you will use; note that phrasing here can make quite a difference. For example, asking ChatGPT

* “edit the following text: …”, versus
* “edit the following abstract of a scientific article: …”, versus
* “edit the following abstract of a scientific article to make it sound like it was written by a really smart person”, versus
* “make this abstract sound like it was written by a native speaker: … “, versus
* “I have written the following abstract for a scientific article. Can you provide some feedback about it? what should I do to make it sound like it was written by a native speaker? …”

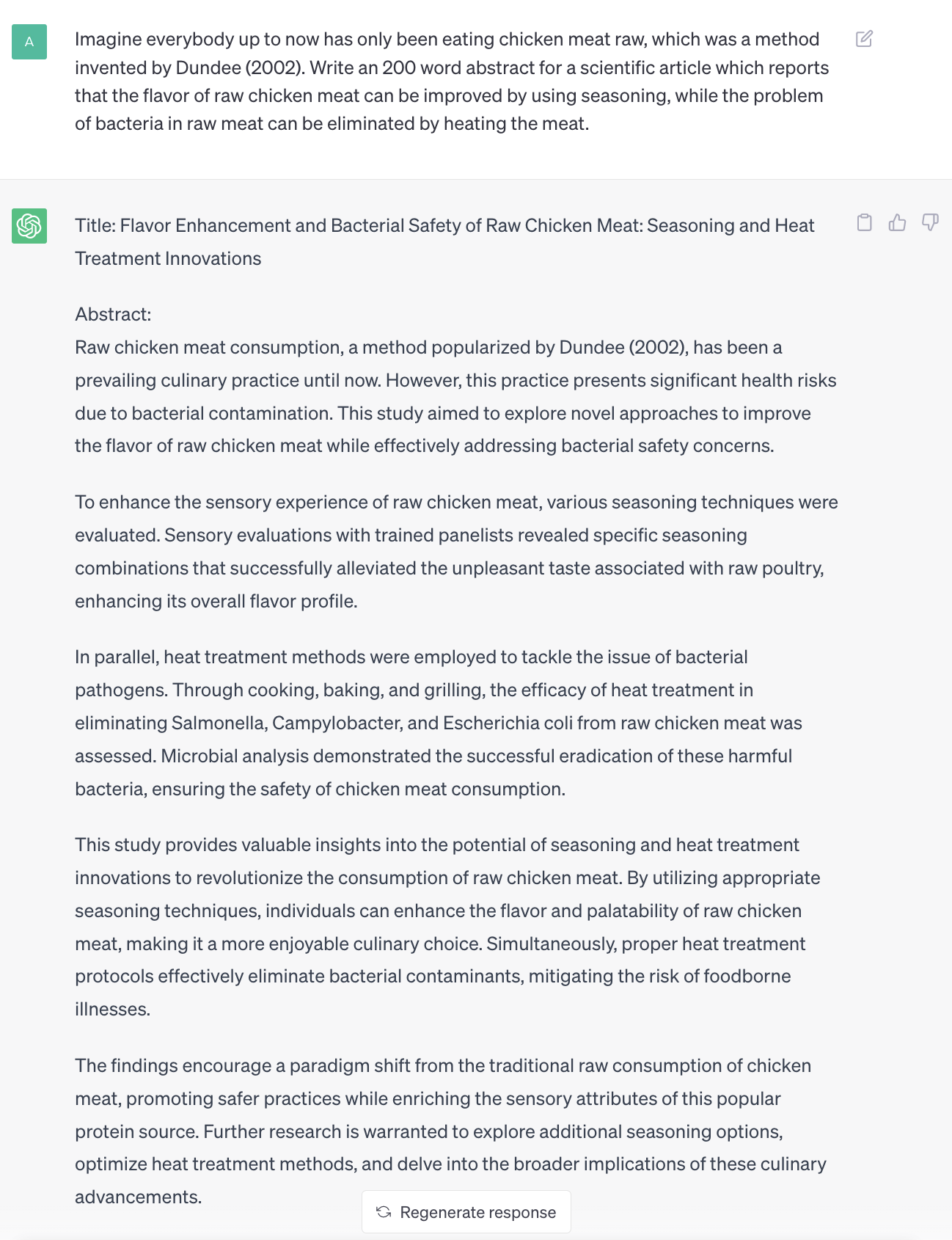
can lead to very different outputs. Similarly, you might get different results if you edit the entire abstract at a time versus focusing on individual sentences.

Then, perform a set of experiments. Keep track of your inputs, the resulting outputs, and your evaluations. Finally, document your experiment with a short write-up (at least 1 page of text, excluding quotes from your AI inputs and outputs). Be sure to include the original abstract, as well as your best modified version of it. The write-up should include:

* Who was in your group and which AI program did you use?
* How did you evaluate text quality?
* Did the AI program help you improve the Chicken midterm abstract?
* If yes, document some best practices for how to improve scientific abstracts – for example, do some types of input prompts work better than others? etc.
* If no, document why using AI did not lead to any improvements. Include examples.

**Submit your group’s short write-up via email to the TA and cc the Professor by noon tomorrow (Friday, 2024.06.07).**

As a word of caution, note that you could also simply ask ChatGPT to write an abstract from scratch. However, in my experience, the resulting output usually has serious shortcomings (example generated back in 2023…):



You should notice several problems:

* First, the abstract is too long, even though the input asks for 200 words.
* Second, the important motivation sentence (“However, this practice presents…”) only mentions the problem of bacteria but skips the problem of flavor.
* Third, the abstract is far too vague. It doesn’t state any concrete results (there are no specific results like “Heat exposure successfully reduced the number of bacteria by X %”), and the motivation is generic (“Raw chicken meat consumption, …, has been a prevailing culinary practice until now.”).
* Finally, the discussion of implications at the end is too long and should be cut.

**Do not hand in any work like this for your final writing assignments!**