

## Project Part 2: Reflection

Annie Chen

1. Consider the manipulation we conducted in the experiment. Would you say this was a good way to test our hypothesis? Explain why or why not.

I would say that it is reasonable to manipulate and provide advice from either ChatGPT or experts to see the influence on people's confidence. However, we did not take into account how people interpret the advice differently, this may lead to the challenges of excludability.

2. Compare your findings to your expectations from the pre-analysis. Did the results turn out the way you expected? Did you find anything surprising in the results or the data? (It's fine if the answer is no.)

The result did not turn out the way I expected. To be more specific, I hypothesize that the advice from ChatGPT and experts will affect people's confidence, but the result is not statistically significant. Moreover, I also stated that people's familiarity with ChatGPT might affect people's confidence in their answers; however, it's totally opposite to my hypothesis.

3. Are there any questions you wished we had asked? Or any ways we could have changed this experiment to improve it? For example perhaps there is an additional subgroup we would have liked to explore, or an additional validity check that would require more data.

I have two ideas for modifying the experiment, and I also have a question regarding the experiment.

First, I would like to first ask about people's confidence in their answers before showing any advice either from ChatGPT or experts. In this way, we can compare their own confidence level before and after seeing those advice.

Secondly, I think we may change the estimating crowd size task into a more complicated or step-by-step task that we can ensure everyone is following the advice (from ChatGPT or expert) to do the task. In this way, we can avoid the uncertainty that people might not be following the advice or just randomly provide their guesses.

I'm considering another question, but I'm not entirely sure about its practicality: If we were to show participants how ChatGPT uses its method to estimate a crowd size and provide a specific number, would ChatGPT's direct answer affect people's

confidence in their own estimates? (cuz I used the photo and asked ChatGPT about the estimated crowd size, it replied me with a number of 7000, which is far away from the exact number of the crowd. I'm thinking that is it ethical or soundable to show the answer of the crowd size and ask people about their answer?)

4. Re-assess your own submission. Describe your proposal in a few sentences. Having seen how this experiment played out, how would you have changed your experiment?

I am really honored to be selected from the majority of our fellow students. I think this topic on how ChatGPT affects people's thinking or confidence is truly an interesting topic to conduct experiments on. However, I think this experiment still has a lot of room for improvement. For example, we should consider the nuances in how advice is received and interpreted by the participants. I also want to hear how our classmates think of this experiment and want to know their feedback.

I am also a bit lost on how to include the qualitative data (strategy) to do further enrichment on the understanding of the results.