The question our team is trying to research is whether gender bias impacts the perception of an individual’s level of aggressiveness. In terms of the experiment design, participants will be asked to read an article describing a typical workplace scenario. The treatment group will be given the story with the main character as a woman. This is achievable by giving an obvious feminine name (i.e. Jenny, Annie, Elizabeth). The control group will be given the exact same story except the name of the main character is changed to an obvious masculine name (i.e. Jack, Daniel, William). In both cases, we will attempt to choose names with neutral sentiments. At the end of the article, the experiment subjects will need to answer the question: On a scale of 1-10, how aggressive do you feel (the main character) was? To conduct this experiment, we will be recruiting subjects through Amazon Mechanical Turk (MTurk). The story and the survey will be hosted on another platform since Mechanical Turk does not provide enough experiment design power especially for randomization. We are looking into SurveyMonkey and Qualtrics to determine a more suitable platform for our purpose.

Currently we have the story created. We are also in the process of figuring out the survey logistics. One of the issues we addressed in previous meetings is non-compliance, as users may simply answer the questions without reading the story. Knowing that there will be many ways people would “cheat” on the internet, there is no perfect solution for this. One possible way to test for compliance is by tracking the length of time each subject will spend to finish the survey.

Although we have a basic structure of the experiment, the details are yet to be fleshed out. We need to figure out the linkage between MTurk and our survey platform, making sure that the transition is done right so that the participants would be able to navigate to the survey easily. Also, we need to make sure our randomization is done right. It is yet to be discussed how we will assign participants to the treatment and control group in the survey platform.

There are some components of the experiment we are still undecided about and would like to hear advice/feedback from peers and professors. First, we pointed out that we would like to capture demographic information. We’ve also considered blocking on demographics. But we are unsure of how much blocking we should do or if we need any blocking for our experiment. Second, we are undetermined on the price point per each task to pay for participants. It would be helpful to know the average price per task for the past experiments that decided to use MTurk for subjects recruitment. Third, we considered calculating the statistical power before running the experiment. However at this point, we believe we have insufficient information to do so. Fourth, currently we are only concerned about controlling the gender of the main character for this experiment. We noticed that the gender of the second character may also influence how people perceive the main character. However, we are unsure whether we should keep the variation in one character or both.

It is worth noting that our experiment relies on off the shelf solutions, so there will be no software needed to be written. In order to ensure that we have the resources necessary to generate a sample size sufficient to detect a statistically significant effect, we’d like to request the maximum amount of $500 to fund our usage of MTurk, which will follow in a separate email. We would be more likely to recruit a larger group of participants with more grants, and hence maximize statistical power, accounting for expected sample attrition from users not engaging in experiment as intended.