**Essay 2: Experimental Proposal**

**Experiments and Casualities**

**Younus Ahmed**

**09/25/2018**

How likely are people to conform to opinions of a group or other viewpoints?

The outcome of the following question has many implications for marketing, politics, social media, etc. In [a study](http://www.nytimes.com/2014/11/21/upshot/social-media-deepens-partisan-divides-but-not-always.html) that examined 2.2 million politically engaged Twitter users who followed at least one congressional candidate during the 2012 elections, researchers found that 90 percent of candidate tweets seen by Democrat voters came from Democrats, while 90 percent of candidate tweets seen by Republican voters came from Republicans. In this sense, social media helps create that much-talked-about echo chamber of opinion. If the outcome of the study can identify that humans have a subconscious proclivity towards following the herd, even for most trivial biases one may have towards current events, perhaps one can feel a greater sense of urgency to question their own biases before picking a side on controversial topics such as free universal health care or the motives behind the actions of our current President.

In the 1950’s, a psychologist named Solomon Asch conducted an experiment on conformity, and found that group pressure can be so powerful that it can warp an individual’s perception on the length of a line. Asch conducted an experiment whereby 50 male students from Swarthmore College in the USA participated in a ‘vision test.’ Using a line judgment task, Asch put a naive participant in a room with seven confederates. The confederates had agreed in advance what their responses would be when presented with the line task.  The real participant did not know this and was led to believe that the other seven participants were also real participants like himself. Each person in the room had to state aloud which line (A, B or C) was most like the longest line. The answer was always obvious.  After the fake participants all stated aloud their blatantly incorrect answers, the real participant sat at the end of the row and gave his or her answer last. After 12 trials, a staggering 75% of participants conformed at least once.

Unfortunately, the experiment that Solomon Asch conducted led to an inherently biased set of outcomes. The selected subjects were all male subjects from within the same age group. He did not vary the demographic of the subjects in the study. They were not selected randomly. Additionally, it does not seem that people are generally faced with the issue of choosing a relationship between two-line lengths. The outcome does not seem very applicable to everyday people. Solomon’s experiment also did not seem to obtain a test statistic between treatment and control to determine whether or not the results he obtained were statistically significant. He simply took a point estimate and arrived at his conclusion with it.

The proposed experiment is a modified version of Solomon’s experiment that attempts to incorporate a valid experimental setup that leads to an unbiased estimate. The proposed experiment requires each of its participants to take an auditory survey of approximately 20 questions that addresses a controversial topic that permeates our media today. The reason that I chose to use an auditory survey versus the typical text survey is because studies like the one in “The influence of auditory cues on the perception of, and responses to, food and drink” by Charles Spence have shown the large influence that auditory cues have on perceptions. This will be more relevant later on as I outline the details of the treatment in this experiment. Each of the participants will be presented with four different multiple-choice answers to each of the questions in the survey.

The setup of this survey is quite unique from the usual click-and-submit type survey style. After the participant is presented with a question from the audio device, he or she will have to submit their response verbally. The survey will require that each of the participants have a microphone that they can use to submit their answer. However, before the participant has a chance to answer the question, participants in the treatment group will be required to listen to 10 other “participant’s” answers to each question with a brief explanation for why they chose that answer. The 10 fake participants will have consistently chosen the same answer with a slightly modified explanation of why they chose it. In order to ensure that the participant does not have sufficient amount of time to think of an answer before they hear the other participants, the 10 fake answers will be presented right after the question and multiple-choice answers are presented in audio format. The participant will then have the chance to answer the question. To further the effect of the illusion, the participant will have the option of “sharing” his audio answer to the next test subject who may take the survey. This is an obvious misdirection since every subject in the treatment group will hear the same 10 fake participants over and over again. On the other hand, participants in the control group would answer each of the audio questions without any influence from outside pressure. They would simply her each question with the possible answer choices and they will be asked to provide their own answers.

The test statistic that will be gathered from each participant from both groups would be the number of answers that were chosen that were identical to those that were chosen by the fake participants. The same statistic would apply to the control group even though they do not hear the fake answers. Although one may criticize this measurement, it would be highly unlikely that a participant in the control group would consistently select the same answer for every question as the fake answer chosen by the fake participants. In the case where the participant did answer every question with the same answer as the fake participants, their test statistic would be averaged out by the outcomes of the other participants in the control group.

In order to minimize the bias associated with the specific topic that the survey will be covering, there will be two sets of surveys that cover two different controversial topics. Half of the subjects will be randomly chosen for one block to take the first survey while the other half get assigned to the second block to take the other survey. A separate randomization process will be implemented to assign individuals within the subgroups to treatment and control groups with a minimum of 20% of participants belonging one of two groups. Because the blocking variable that was chosen was the topic at hand, and it is believed that the topic of the survey is a covariate that should be accounted for, the standard error of the average treatment effect should be reduced.

Subjects can be retrieved from many sources for this experiment. If the subject is not aware of the nature of the surveys that are administered to them, then the potential for selection bias is minimized. Participants can be obtained from Facebook or other forms of social media, Mechanical Turks, and friends or family. In order to truly gain insight from a true randomization process, Mechanical Turks would be the most optimal medium through which to obtain subjects. It could be possible that subjects obtained from Facebook are already inclined to be influenced by peer pressure, given the rate of controversial claims that are made on Facebook every minute. It would also narrow the demographic to an age range that is more prone to using the internet, which could lead to a biased outcome.

Once data is collected from all the subjects from both the treatment and control group, since our sample is not very large, a randomization inference study on each block would be appropriate to use under the sharp null hypothesis. The null hypothesis is that there is no difference in outcome between the control group and treatment group. This implies there is no significant impact from outside pressure on the participants ability to answer a question that is representative of his or her own values. If the null hypothesis is unlikely, then in order to answer the question at hand, we would expect that the average potential outcome of treatment group would be higher than the average potential outcome of the control group for both blocks. The reason for this is that a higher test statistic implies that the participant answered more of the questions identically to the answers of the fake participants. As a result, a one-sided test will be conducted for the hypothesis testing. After the average treatment effect is collected from both blocks and weighted by the participants in each block at the end of the experiment, it will be measured against the distribution of ATE’s generated by assuming the sharp null hypothesis. A p-value of 0.05 will be used as a measure of statistical significance. Additionally, in order to get an understanding of statistical power, the random inference simulation will be conducted 50,000 times to get a p-value distribution.

There could be interesting covariates that can be collected from this experiment. If the null hypothesis can be rejected in favor of the alternate hypothesis, one can measure if either age or gender has any impact on one’s propensity to conform to others. A regression can be created before beginning the experiment, after collecting statistics on the participants, and then covariates like age,race, and gender can be regressed on.