Essay 2: Experimental Research Proposal “Personalization to increase engagement”

W241: Field Experiments

One of the most effective mechanisms for organizations and individuals to grow and improve is through acting on feedback. In my proposal I want to explore the relationship between personalized attention and feedback response rates and quality. More concretely, the question that I want to ask is “do more personalized requests for feedback increase the quality and quantity of feedback responses from subjects?”. This question is within the context of widely disseminated feedback requests to multiple people for a single event within an organization (e.g. a company’s quarterly “all hands” meeting).

I find this question interesting because it attempts to determine whether more personalized feedback requests can increase feedback request success rates. I believe that increasing feedback request success rates is critical for companies to get visibility into how their personnel think and feel about the company, which can help the company gain better understanding about itself and improve itself.

For several companies and organizations, one of the most valuable assets they have is their personnel. Better understanding personnel and identifying trends in how they feel and think about the organization in a timely and effective manner can be critical for a company to maintain personnel metrics such as morale and retention.

If it is the case that more personalized feedback requests cause an increase in (1) responders’ engagement or (2) quality of responders’ engagement in said feedback requests then companies can use this knowledge to better procure valuable insights on its personnel. This knowledge can help companies systematically evaluate and develop company culture as well as improve morale, performance and retention.

The subjects for this experiment would be individuals from the company that I currently work for. Given that the company hosts multiple “All Hands” meetings as well as other meetings where feedback is requested I can tap into the pool of individuals that attend meetings where feedback is subsequently requested as subjects for the experiment.

The treatment that subjects will be exposed to is a more personalized feedback request after attending a meeting: a feedback request with customized language that personally addresses the recipient and ‘specially’ asks for feedback. All attendants to the meeting would receive a feedback request after the meeting

but these requests would differ for the control and treatment groups respectively. In the case of subjects in the control group a generic feedback request would be sent, whereas in the case of subjects in the treatment group a personalized feedback request would be sent out.

Another potential layer to the experiment could be an additional treatment to meeting attendees that have not responded to the feedback request within a period of time. In this case the treatment could be a personalized email to an attendee that has not responded to the feedback form asking him or her to fill out the form. The control group could simply not receive an email prompting a response or alternatively could receive an impersonal reminder to submit a feedback response.

Subjects in the experiment would be placed in treatment through random assignment. For all the attendants of a set of meetings occurring at the company headquarters for a set period of time that warrant a feedback request being sent, confirmed attendees would be randomly assigned to treatment or control.

I plan to collect gender, historical feedback responses and position as covariates. I suspect that whether an individual submits a feedback response (and whether it is high quality or not) may be a function of the subject's gender, the subject's previous responsiveness to feedback requests as well as the subject's place in the company's hierarchy. Since I believe that these parameters inform the outcomes that I am measuring (quantity and quality of successful feedback requests), I believe that including these covariates will increase the precision with which effects can be measured.

Since I plan to use covariates in the study, I intend to use blocking in order to detect effects with a higher precision. Blocking would be implemented by block random assignment of subjects to treatment based on whether they are male or female, as well as whether they are individual contributors or managers of personnel.

The outcome measure that I care about the most is whether the subject submitted feedback. That is, in aggregate I want to measure the percentage of total feedback requests that were successfully filled out and returned by the subjects. This would be a binary indicator for individual subjects, but the average outcome would be the ratio of total feedback responses received to total feedback requests issued.

A secondary outcome that I plan on collecting is the quality of the feedback responses: that is, how much of the feedback form was completed for successful responses. This would be measured as the percentage of questions answered in the feedback form as a function of the total questions asked in the feedback request. An alternative outcome that could be measured to determine the quality

of feedback responses is the number of words used to answer non-multiple choice questions in the feedback responses.

Once the data is collected I plan on conducting statistical analysis to determine the estimated average treatment effect as well as determine through randomization inference whether the results obtained are likely to occur under the sharp null hypothesis that personalizing feedback requests do not affect each subject's individual propensity to respond to a feedback request (or the quality of the feedback responses).

First I would calculate the estimated average treatment effect for each of the four blocks resulting from the covariates mentioned earlier using a difference in means estimator. Then I would do block randomization inference to determine how likely I would have been to obtain the estimated average treatment effect by chance under the sharp null hypothesis. That is, obtain a p-value from the point estimate and compare it against a 5% critical level. Finally, I would construct a 95% confidence interval based on the estimate obtained from the experiment. It is important to note that the global estimated average treatment effect as well as its associated standard error would have to be computed taking into consideration the estimated average treatment effect from each of the blocks in the experiment.

An advantage of this field experiment is that I could easily conduct a pilot study as a dry run for the larger experiment early on in the project. To do so I could simply conduct the experiment on a small set of meetings with feedback requests sent after the meetings. Another advantage of this proposal is that the infrastructure to conduct the experiment is already established at the company. Feedback requests are routinely emailed to attendees after meetings through specialized software for constructing feedback forms and analyzing feedback responses. This would allow the team conducting the experiment to focus on the experimental design and analysis rather than on issues of data collection and outcome measurement.

Finally, the largest risk that I see with the experiment is getting buy-in from meeting organizers to conduct the experiment. Since feedback responses give an important glimpse into the performance of a private company, the meeting organizers may not want to risk inferences from the feedback responses becoming public, presenting a challenge to their agreement in conducting the experiment. I also believe that abiding with non-interference will be challenging as subjects may hear of the experiment being conducted in the company which could bias their response to being placed in treatment or control. It is conceivable that interference could downwardly bias the estimated average treatment effect as subjects in control that have heard of the experiment being conducted will

respond to feedback requests at a higher rate/with higher quality responses (e.g. potential outcomes to control would be artificially higher).