Impact of Gender on Customer Satisfaction

Yuchen Zhang, Dili Wang, Miguel Jaime, Cassie Seney

April 2019

Within corporate America there is increasingly a focus on improving the diversity of employees, including gender diversity. Research suggests that gender diversity enables better problem solving and can improve the profitability of corporations[1]. Employees of different genders often have different viewpoints, ideas, and insights, and bringing these together can be advantageous for businesses.

In the technology sector, where women are typically underrepresented, studies show that women often need to provide more evidence of their competence than their male counterparts. Women are also more likely to have their judgement questioned in their area of expertise and may be mistaken for someone at a more junior level (2). Other research suggests that women are more likely to be perceived as empathic and caring (3).

Technology companies that provide customer service and support often use customer surveys to obtain ratings of the level of support provided by their employees. These ratings reflect upon the success of the support organization as a whole and also on the abilities of the individual technical support engineers, whose career advancement may be impacted by the survey results.

This experiment aims to determine whether the gender of a technical customer service provider impacts received customer service ratings. In a technologically complex environment where specialized skills and experience are required to provide a service, do customers value the work performed by men and women equally? If not, how do customer perceptions of value provided differ based on the gender of the service provider? Does gender matter more for bad support than good?

If a difference is found in the customer service ratings based on gender that biases against women, then corporations need to take this into account when considering the career advancement of women in the field. Not doing so may guide corporations into unintentionally favoring one gender over another in career advancement instead of promoting service providers (and possibly all employees) based on merit. This, in turn, may impact corporate diversity and fairness goals.

Experiment Design

Hypothesis Testing

The methodology we used to investigate the effect of technical support representative gender on customer satisfaction was via online survey. Our outcome survey questions were inspired by customer satisfaction survey templates in several online survey tools. In a technical support field, we were particularly interested in knowing whether perception of the knowledge of the support representative varied with their gender. Finally, we were also interested in ascertaining whether a level of personal commitment would in any way alter our subjects gender based evaluation of customer support.

Our survey questions asked participants to provide overall customer satisfaction rating, more specific ratings for the representative’s understanding and knowledge, whether they would be personally willing to receive support from this technical support representative, and to specify a monetary gift to be offered to the support representative based on the level of support received / satisfaction.

For each survey question, we tested four hypotheses:

- Female support representatives receive worse ratings than males
- This discrepancy is greater when agents provide bad support compared with when agents provide good support
- Gender based perceptions of customer support vary depending on the gender of the subject
- Gender based perceptions of customer support vary depending on the age of the subject

Online Experiment

In order to evaluate the potential causal relationship between gender and customer support ratings, we designed an experiment with a treatment applied to randomly selected treatment and control groups. Random assignment of treatment allows us to make apples to apples comparisons between two populations, since we know they are equivalent. When the only variation between the groups is the treatment of interest, we can more reliably estimate causal effects.

To apply treatment or control consistently across our subjects we used written transcripts of a support interaction delivered online. This virtual written approach allows the gender of the customer support representative to be artificially simulated. It also eliminates other sources of variation that would typically occur with research conducted in person, since live actors vary across many dimensions.

Subjects were sourced online through a survey service providing randomly selected participants meeting the blocking requirements of our design. Each experimental subject was presented with a written sample of online chat interactions between a customer service representative and a customer requesting support. After reading the chat dialog the subject was asked to evaluate the level of support provided via an administered online survey.

Multifactor Design

Our experiment used a 2x2 factorial design with two gender treatments and two support scenario treatments. Subjects were allocated randomly to each gender-scenario combination.

Gender Treatment

Subjects in the control group were shown support interactions where the support representative was “male”. Subjects in the treatment group were shown identical scenarios where the support representative was “female”. In this research, we are not addressing more complex LGBT gender scenarios, which other studies have shown to involve potential bias in the workplace. For this initial study, we have simplified to two primary gender cases: male and female.

Initially, the gender of the representative was signaled in three ways to ensure that the treatment was effective:

1. The name of each representative as unambiguously male or female (“Angela” and “Henry”)
2. Text color signaling gender (pink and blue)
3. Race-neutral gender emojis

After the pilot study we also incorporated gender references in the survey questions.

Demand Effect

In our experiment, we attempted to conceal the connection between the survey and the gender treatment to avoid the risk of subjects tailoring their responses to what they believed we wanted to hear. While we want our subjects to notice the gender of the support representative, at the same time we do not want them to consciously tailor their survey responses to that factor. We were careful to avoid asking survey questions which may hint at the purpose of the survey.

Scenario Treatment

We included two customer support scenarios demonstrating different support agent capabilities to resolve the reported problem. The first scenario showed a successful resolution of the problem allowing us to measure how different genders are praised for performing the same technical work. The second scenario was a less satisfactory resolution, to provide more opportunity for critical analysis of the representative's performance.

While the scenarios dealt with complex technology, they were phrased at a level that could be understood by a layperson. The scenario involved a technology and situation we expect would be familiar and relatable for most subjects. Scenarios were designed to be short, requiring only a couple of minutes reading time.

Copies of the scripts are provided in the Appendix to this document.

Covariates and Blocking

Gender and Age

Research suggests that the gender and age of the subject may impact attitudes towards the gender of the support representative. These variables may then be related to the outcome of our gender treatment. The implementation of our survey keeps our subject selection balanced on gender and age.

Predisposition Toward Customer Support

Randomized selection is expected to provide a pool of test subjects with a variety of different prior experiences with technical support agents. To capture information about this characteristic for balance checks and blocking, we included pre-treatment questions to capture the subjects' most recent experience and general attitude toward customer support. We used blocking to restrict ourselves to randomizations that keep treatment and control groups similar for this variable.

Good and Bad Controls

Questions about our subjects' previous support experiences were asked prior to the survey to ensure that we captured covariate data that was unaffected by the treatment. Capturing this data after the survey could render them bad controls, as they could be impacted by the effect of the survey treatment.

Evolution of our Design

Our design went through several iterations to reach the final survey form.

Gender Identification Considerations

Our original design used the name of the support representative to identify gender. This may not be sufficiently compelling to produce a treatment effect. We considered a few options to improve this design:

- Including language that might signal gender. Certain online articles suggest that women use certain words and phrases more than men, to their detriment. We considered modifying the text of the female scenarios by incorporating these phrases. We decided against this approach since our treatment would then be based on the type of language used, rather than simply on gender as originally intended.
- Using audio recording to clearly identify the gender of the support representative. This design would introduce other factors, such as the race, age, and attitude of the support representative, which might be deduced from the audio recording. Our text-based approach was designed to avoid this issue.
- Using text color and emojis. We incorporated these into our text-based approach as they allow us to emphasize gender without introducing other factors.

Our final design used unambiguous, race-neutral names ("Angela" and "Henry") for the support representatives, text color signaling gender (pink and blue), and race-neutral gender emojis to highlight support representative gender.

Non-Interference and Spillover Effect Considerations

In our original design, we presented both the good and bad support scenarios to the same experimental subject. This design would need to satisfy the additional non-interference assumptions of a same-subject experiment. When the same subject is exposed to two different treatment effects, such as the scenarios in our experiment, there are spillover effects to be considered.

To satisfy the "no persistence" assumption, we would need to assume that exposure to a good support scenario would have no effect on the rating applied to the subsequent bad one or vice versa. In cases where the gender was varied, we could also violate the assumption for same subject experiments on the gender treatment. If we could not guarantee "no persistence", our analysis of outcomes would need to focus on the first responses for each of the subjects to address persistence concerns. The subsequent response would be considered a cumulative effect of the first response followed by the second.

The "no anticipation" assumption requires that expectation of future treatment has no effect. For that assumption to be met in this experiment, subjects must provide responses to the first scenario, which are not impacted by knowledge of the second scenario.

In our final design, we chose to present one scenario and gender treatment to each randomly selected subject to avoid non-interference concerns.

Pilot Study

The objectives of our pilot study were to obtain feedback on our design and to perform statistical analysis to evaluate our experiment.

To recruit subjects for the pilot study, we each nominated friends and family who were willing to participate. This provided us with a pool of 32 subjects with an even distribution of gender. We also had a good spread of ages, although this was not a balanced distribution. For the pilot study we were more interested in obtaining a wide range of perspectives than in blocking balance.

With this small group of subjects, we assigned most of the subjects to evaluate the two bad support scenarios. These results were used for our data analysis. We selected four subjects to review the good scenarios, and then randomly assigned the remaining 28 subjects to Male or Female bad scenarios.

The table below shows the distribution of gender and ages for our randomly assigned pilot study group:

Table 0:

Gender		Age	
F	M	≤35	>35
15	13	19	9

For each scenario, we had a separate Google form of the survey questions. We sent emails or text messages to each of the subjects with the survey link included.

Experiment Design Evaluation

Predisposition Toward Customer Support

We included four pre-treatment questions designed to measure the subject's general disposition towards customer support:

- *Recall your most recent interaction with a customer support agent. This interaction could be on the phone, via online chat, or in person. How would you rate your experience?*
- *Have you ever experienced support that you would rate as Outstanding?*
- *Rate this statement: contacting customer support when I need to solve a problem is a good use of my time.*
- *How many times a month do you contact customer support?*

We intended to use just one of these questions for blocking purposes in our survey, and use data from the pilot study to select the most appropriate question.

Attention

We included a question to verify that the subject had actually read the transcript.

- *Which system was the customer experiencing a problem using?*

Within the pilot study this question could also help determine whether subjects were able to understand the scenario.

Gender Treatment

Our treatment effect requires that the subject notice the gender of the customer support representative. We asked this question in our pilot study. We also checked to see whether our respondents were aware of consciously changing their responses based on the gender of the support representative. These were our included questions:

- *Did you notice the support agent's gender?*
- *If you answered yes in the question above, do you feel, in retrospect, that the gender of the support agent influenced your rating of the support agent in anyway?*

Design Feedback

We also asked questions to solicit freeform feedback from our study participants on our survey design.

- *Did you find any of the survey questions confusing or strange? If so, which ones?*
- *Did you find the view of the customer service script, which consisted of 2 columns, confusing?*
- *Do you have any other feedback or questions you have for us?*

Pilot Study Data Analysis

Data analysis of the pilot study focussed on correlations and basic modelling of treatment effects. With a small sample size sourced from friends and family we had concerns that the sample size would not be large enough to wash out individual variations. Nevertheless, our covariate balance check passed and we were able to derive some useful conclusions.

Pre-treatment Questions

Of our four pilot study pre-treatment questions we selected one to use in the survey as a blocking variable measuring predisposition to customer support:

- *Recall your most recent interaction with a technical customer support agent. This interaction could be on the phone, via online chat, or in person. How would you rate your experience?*

Correlation analysis of the pilot study data revealed that this pre-treatment question had the highest correlation with outcome variables and could therefore do a better job of explaining variance of outcome when used for blocking. The frequency histogram for this variable showed a good distribution across the range of values even within our small set of respondents. It also appeared to be easier and less confusing to answer based on respondent feedback.

The pre-treatment question will also be changed to use the phrase “technical customer support” in response to feedback from a pilot study subject who would have responded differently if the question had been more specific to the type of support.

We noticed a negative correlation between our outcome variables and the pre-treatment question which asked whether the subject had ever experienced outstanding customer support. We were concerned that recollection of previous outstanding support may have caused the subjects to respond more critically in a survey where the customer support was suboptimal. We decided against including this question as a precaution to guard against persistence bias.

Outcome Variables

We observed two outliers of large donations of money for poor support. Most respondents who agreed to a gift selected our lowest amount. We decided to change the scale of our donations to smaller amounts.

Analysis of correlations between outcome variables found a weaker correlation between the gift card question and other outcome variables measuring customer satisfaction. For the 32 pilot respondents who were assigned the bad support scenario scripts, we decided to perform regression analysis on the survey results for the 5 outcome variables. While we did not find a statistically significant treatment effect amongst the outcome variables, all of our estimated treatment effects were negative except for one. We also found that gift card was the only outcome variable with a positive correlation to our gender treatment variable, meaning that although female representatives tended to rate lower on other customer satisfaction metrics they were more likely to receive a gift card. We felt that if this trend were observed in the actual survey, it would be an interesting finding.

We also used the robust standard errors obtained from the linear models to calculate the minimum ideal sample size, or the sample size needed for our standard error to be small enough such that the treatment effect is statistically significant. For each outcome variable, we calculated the ideal sample size under the assumption that the SE is indirectly proportional to the square root of the sample size.

Table 1:

Outcome	Gender treatment effect	Minimum Ideal Sample Size
Overall Customer Support	-0.1743 (0.1675)	100
Support Rep Understanding	-0.1846 (0.1833)	107
Support Rep Knowledge	-0.09743 (0.1577)	282
Request Support Rep Again	-0.03077 (0.1681)	3212
Gift Card	0.2359 (0.1873)	68

The minimum ideal sample sizes calculated here pertain only to the bad support scenario, which would consist of half of the total number of respondents in the actual survey. In order to achieve statistically significant average treatment effects, we would need a total sample size that exceeds twice that of the calculated minimum sample sizes. Given our aim of approximately 250 respondents in the actual experiment, there were some indications from the pilot results that we could potentially detect significant treatment effects in 3 of the 5 outcome variables. However, we were cautious of these results as they may not necessarily be representative of a larger pool of subjects. Unlike our friends and colleagues who share our similar principles, ideals, and even cultural backgrounds, we have much less information on the respondents hired by Lucid, who may not respond to the treatment in the same way as our pilot subjects. Nevertheless, we proceeded to optimize our experimental design based on the pilot results in order to mitigate any potential confusions in the actual survey.

Design Changes Post Pilot Study

Gender Treatment

The pilot study revealed that 40% of our respondents did not notice the gender of the support representative. Modifying the survey questions to incorporate the name and gender pronoun of the support representative within the text of the questions should improve the likelihood of the treatment being effective.

Scenario Feedback

We received the following feedback regarding our scenarios:

- Some respondents indicated that the scenarios were too long
- Some indicated that the bad support scenario did not seem complete as the support ticket was not closed

To address these concerns we decided to make a slight change to the bad support scenario, where the agent proceeded to close the chat sooner and offered to opening an escalation ticket for the customer.

Attention Question

Our question intended to verify that our respondents had actually read the scenario may have been confusing as it required the respondent to understand the difference between WiFi and Ethernet.

We modified the answers to use simpler and more easily distinguished technologies, and modified the question. Subject was now asked to choose the best category for the customer's issue, and the survey provided clearly differentiable categories:

- *Please choose the best category for this technical support interaction.*

Outcome Variables

One of our survey questions measuring customer satisfaction elicited some pilot study feedback:

- *When a call transcript is reviewed, representatives are eligible for an 'outstanding service' gift card reward. Based on agent's performance, what reward level would you recommend?*

Pilot study respondents made the following observations:

- Some respondents felt uncomfortable deciding how much the company should give the support representative. Some also felt that support representatives should just be paid better.
- Some felt it strange that the representative would automatically receive a reward.
- One respondent recommended changing the intervals of the amounts offered.

We reworded this question as follows:

- *If you think [agent name] should be rewarded for providing great service, select one of the incentive options below.*

Online Feedback Follow Up

Some of the responses we received to our pilot study survey highlighted the challenges of obtaining actionable feedback online. One subject responded with the word “Confusing” when asked what additional feedback they could provide. We followed up in person with some of our subjects to clarify responses where we needed more detail.

Survey Tool and Respondent Pool Selection Process

Evaluation of survey tool and respondent pool vendors was based on criteria including: the ability to implement blocking and scenario randomization, ease of use, and cost. We chose Lucid as our respondent pool provider due to their large respondent pool and lower cost per respondent. We selected Qualtrics as our survey tool since it provided the features we needed to implement our survey and was available for free for Berkeley students.

The use of a vendor supplied respondent pool eliminated any concerns about attrition since respondents who did not complete the entire survey were replaced with another randomly selected respondent from the pool, consistent with our blocking requirements. This also prevented us from performing any analysis on the characteristics of the subjects who dropped out to look for any attrition patterns.

The non-persistence restriction was also not a concern for our survey as Lucid ensures that the same subject can not complete the survey twice.

We requested an even distribution of gender and age for our subjects, so we could use those subject attributes as covariates.

Lucid’s role was limited to providing respondents for our survey. Any logic applied to respondents, such as blocking and subsequent random treatment assignment, would need to be implemented in Qualtrics. The different scenario scripts and its differences (such as names, text color, and emojis) needed to be implemented in a single survey as well.

Working with the respondent pool provider was smooth and without incident. Once our survey was ready and implemented with the desired logic, we worked with a Lucid Project Manager to test how our survey would work with their processes. After conducting a handful of tests, we conducted a soft launch with 25 respondents. Once we confirmed the survey and data collection were working well, we approved a full launch. We completed the full 250 respondents within 12 hours of go-live. Lucid handled any attrition and reported that only 9 subjects started the survey without completing it, possibly a sign of the clarity of our survey.

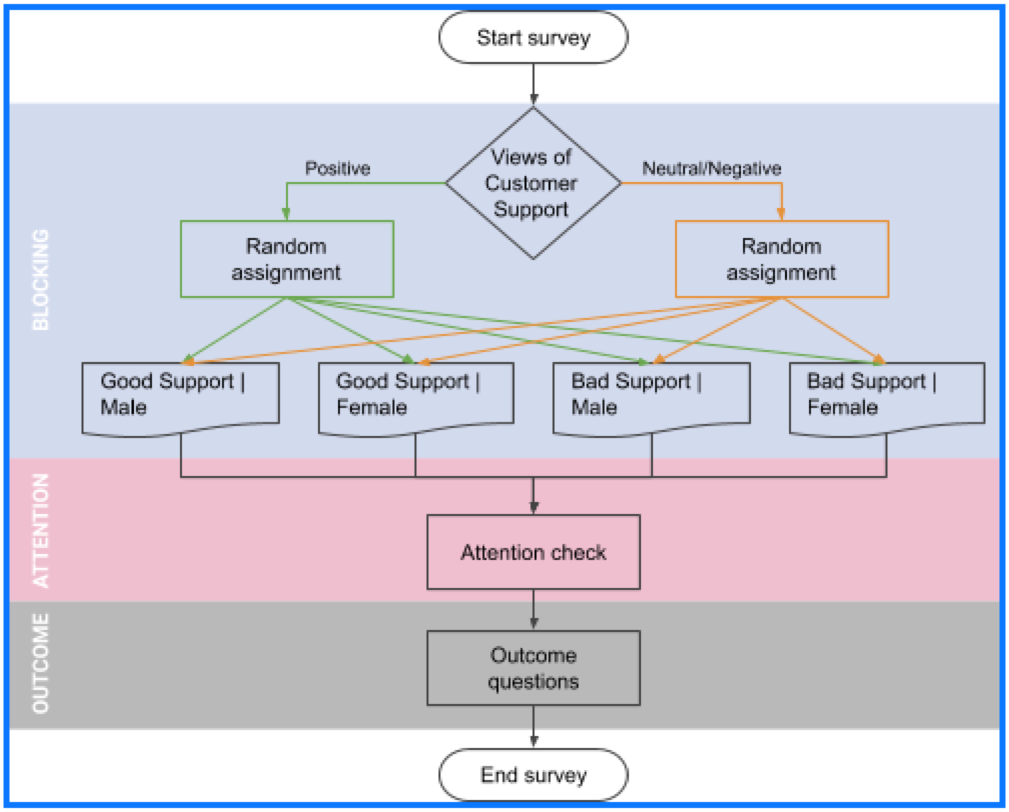
Challenges and Solutions

Implementing Blocking

We originally expected Lucid to handle blocking based on pre-treatment question. This was not possible, since working on the survey was out of scope for Lucid; they simply pass along survey links to their subjects.

That meant we had to implement the logic on the survey side. First, we asked the pre-treatment question. Qualtrics surveys allow for conditional processes, including randomization, based on a subject's response. We blocked subjects based on their pre-treatment question answers, and then each block was randomly assigned to one of the four scripts.

Figure 1:



Mapping Users to Treatments

Once subjects completed reading one of the four scripts, they were then routed back to the same group of questions used to assess attention and potential effect. The challenge though was that Qualtrics was unable to capture which script was shown to which user.

Solving this issue required creating four identical sections with the attention and outcome questions, each with a question identifier that included a prefix indicating the treatment received by the subject. Once all data was collected, we used the prefix to label the scenario and copied the answers to a common set of questions.

Adding Emojis to Transcript

Part of our survey design included signaling gender as much as possible without making it too overt. One of the methods we employed was including emojis in the transcript.

The transcript was created in a Google Doc. When pasting the transcript to the Qualtrics Survey, the tool raised an error: the script was going over the limit of 20,000 characters. This was confusing, since the count in Google showed 1,703 characters for the longest scenario (poor support). The Qualtrics tool stripped out the emojis to get below the character limit.

After performing a number of tests, we discovered the issue: Qualtrics uses HTML for rich text formatting, and the character number counts HTML tags, styling, etc. This would not have been a problem, except that the Google emojis were not imported as native emojis, but as linked images. The "emojis" were resulting in about 5,000 additional characters, since they were added in each line.

We deleted the Google Emojis and inserted the native keyboard emojis in Qualtrics. It was a slight downgrade in aesthetics but solved the problem.

Experiment Results Analysis

Data Loading, Cleaning, and Transformation

Data Loading from Survey Output

Once the data collection process was finalized, we exported the data from Qualtrics. Each row included a referral ID that enabled us to match the respondent with the demographics information provided by Lucid.

As described in the 'Mapping Users to Treatments' section above, our survey asked four sets of identical questions, one for each group (to be able to track which version of the script was being displayed to each respondent). This resulted in each respondent row having 4 sets of outcome questions (for a total of 24 columns, 6 * 4), with only one group filled out:

Figure 2:

BM-C1	BM-O1	BM-O2	BM-O3	BM-O4	BM-O5	BF-C1	BF-O1	BF-O2	BF-O3	BF-O4	BF-O5	GM-C1	GM-O1	GM-O2	GM-O3	GM-O4	GM-O5	GF-C1	GF-O1	GF-O2	GF-O3	GF-O4	GF-O5
						2	2	2	5	3	3												
1	3	5	4	3	2																		
												3	4	1	4	3	2						
												4	3	2	5	$3	4						
																		3	2	2	3	3	4
												5	1	4	3	2	4						
												1	2	2	1	1	2						
																		1	1	3	2	1	2
																		2	1	1	1	1	1
																		2	1	1	2	1	1
												3	4	2	2	2	2						
												2	1	2	1	$1	2						

In order to solve this issue, we transformed each scenario into a table that only includes the questions with the relevant prefix for support quality (bad/good) and agent gender (male/female), i.e. BM, BF, GM, GF. Then we appended the resulting four tables together, and exported it as CSV, which was then loaded into R to conduct our analysis.

Respondent Data

Lucid provided the age and gender of our survey respondents. We converted the gender into a binary value to be used as a covariate. We created a binary subject age value by splitting our subjects into 2 age groups using age 40 as the cutoff.

Likert Scale Conversion

Our preferred outcome variables are binary scores showing whether the female support representative scored better or worse than the male representative in overall customer satisfaction, knowledge and understanding.

Providing a binary approval rating is challenging for survey respondents. Evidence of this was provided in our pilot study feedback when subjects requested a scale rather than a yes/no answer when evaluating support. To accommodate our subjects we used a Likert scale with five ratings for our survey questions. During data transformation we dichotomized the scale by grouping 'Very Negative', 'Negative' and 'Neutral' as group 'Negative or Neutral,' and 'Positive' and 'Very Positive' as group 'Positive'. Discarding the additional details of the rating scale allows us to focus on our desired binary comparison.

Summary Tables, Correlations and Histograms

Data Distribution

Table 2:

Random treatment assignment per recent support experience block.		Treatment			
		Male Good	Female Good	Male Bad	Female Bad
Recent Support Experience	Negative	23	21	13	24
	Positive	45	51	48	41

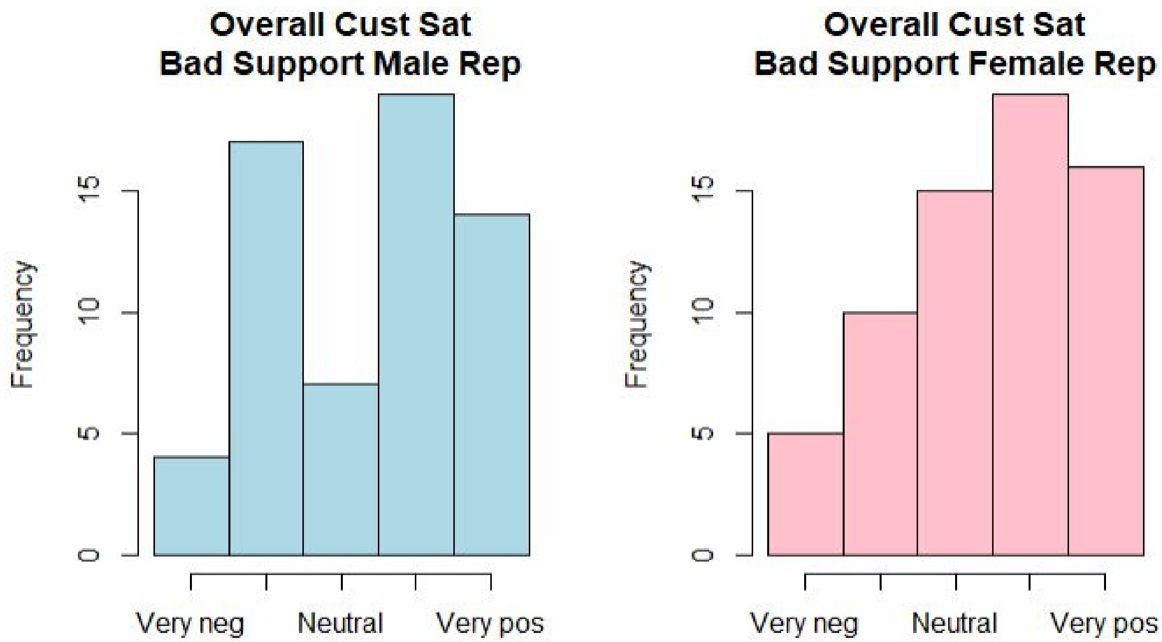
Table 3:

Subject Age		Subject Gender	
<40	≥40	M	F
117	149	132	134

Key Histograms

Histograms were plotted on the raw Likert scale data to examine the distribution of the outcome variables and the relationships between the variables prior to transformation.

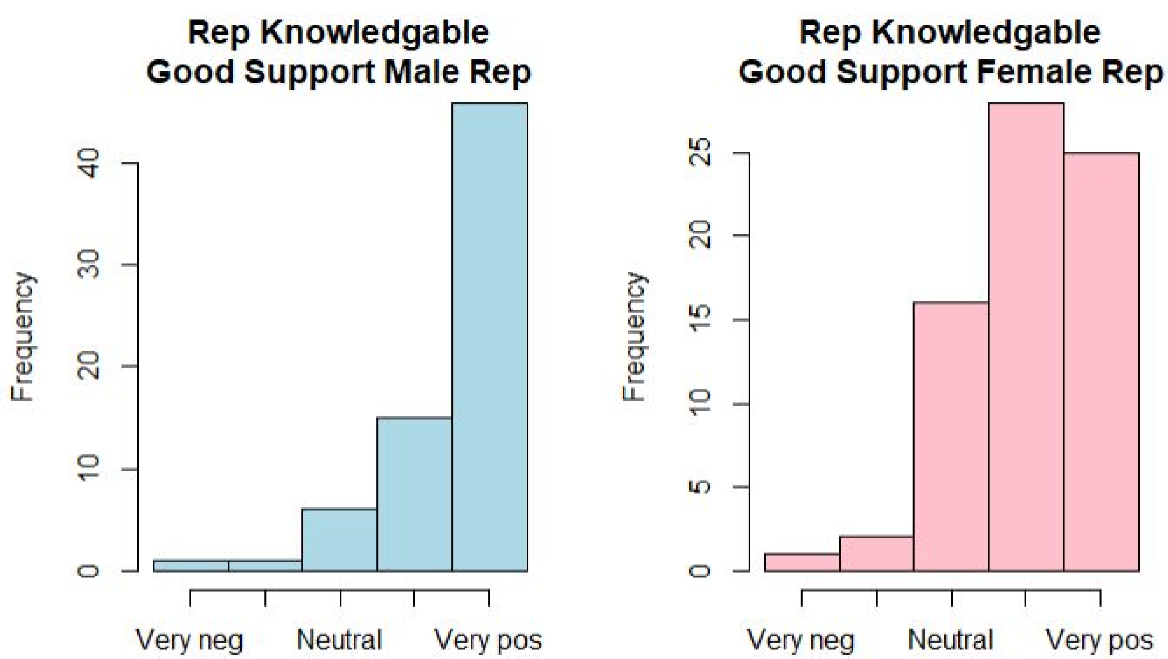
Figure 3:



The data distribution did not always provide support for our hypotheses. The pair of graphs in Figure 3 show the results of the overall customer satisfaction rating for male and female support representatives providing bad support. Based on this distribution, it does not appear that female support representatives are rated more severely for poor support than males.

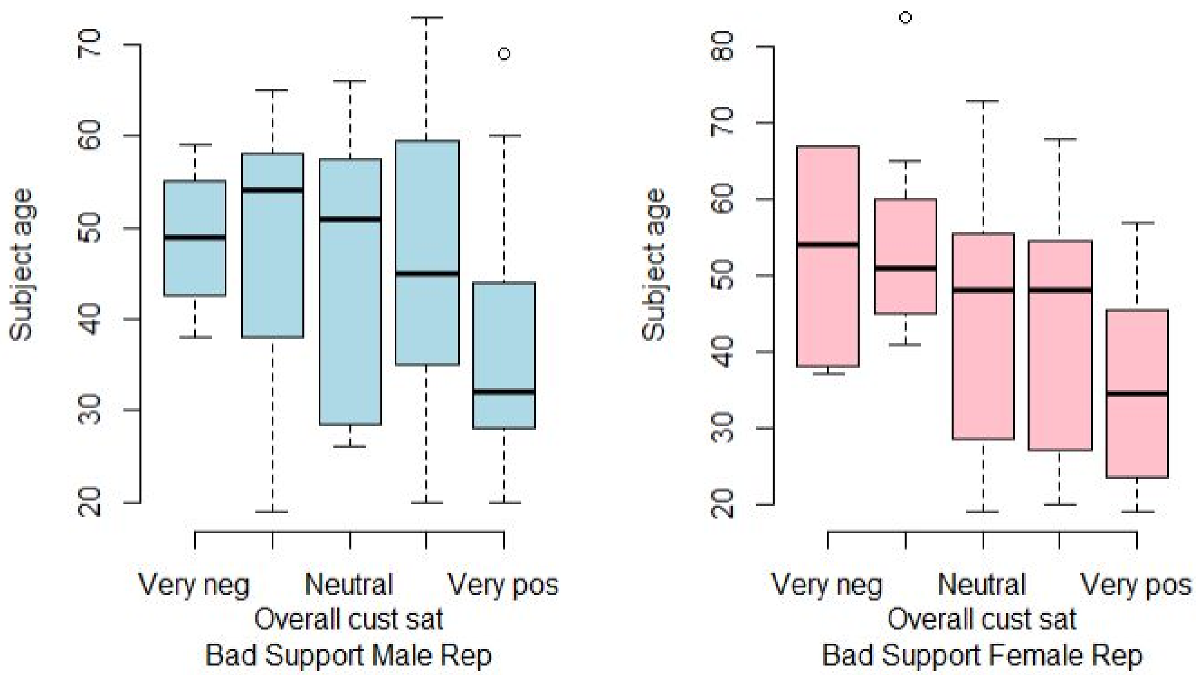
On the other hand, the pair of graphs in Figure 4 show ratings which indicate whether the support representative was knowledgeable for the good support scenario. These histograms show that male support representatives received more positive ratings for being knowledgeable than female representatives providing the same support (note that the y axes are different).

Figure 4:



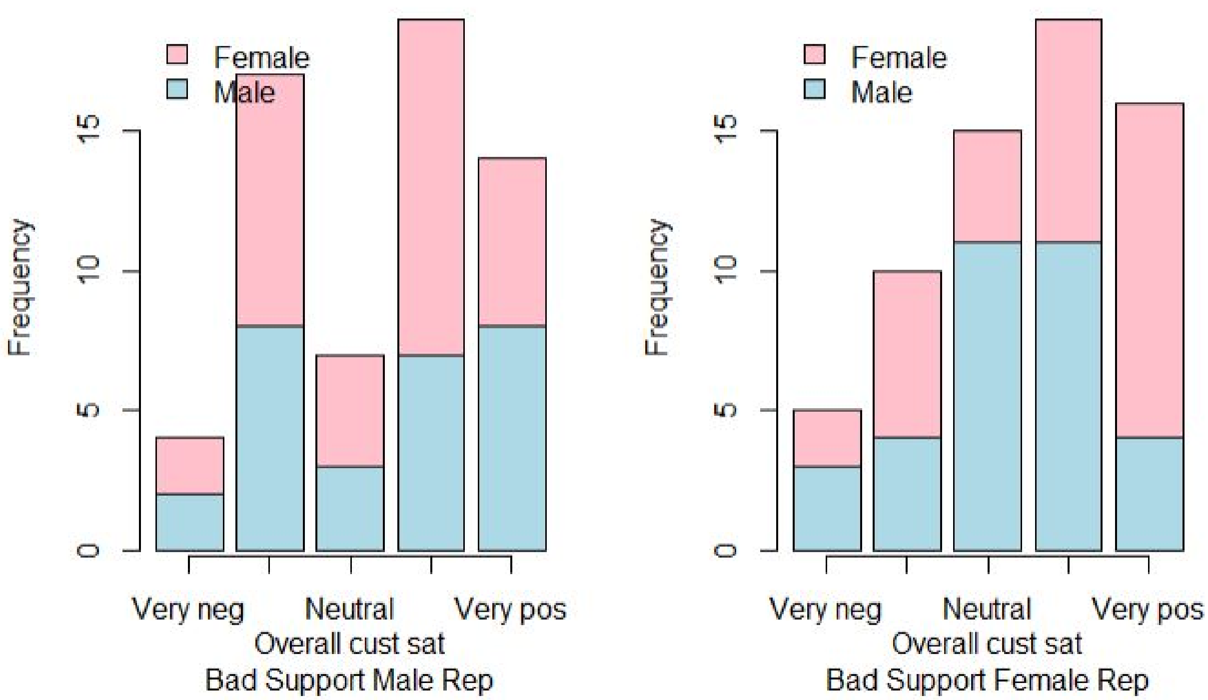
Box plots of subject age shown in Figure 5 illustrate a possible relationship between subject age and lower ratings for female support representatives when support quality is poor. The graphs in Figure 5 suggest that ratings for overall customer satisfaction decrease as subject age increases for female representatives, but this same trend is not evident for male representatives.

Figure 5:



The histograms in Figure 6 show the proportion of male and female survey respondents who provided overall customer satisfaction ratings in the poor support scenario. While it is difficult to see a trend for male support representatives, there does appear to be a higher proportion of female survey respondents who provided high customer satisfaction ratings to female support representatives when support was poor.

Figure 6:



A full set of histograms for all the variables is included in the survey data analysis R notebook.

Correlations

We examined the correlation between independent variables, and with the binary outcome variables. The chi-squared test produces a p-value under the null hypothesis that the variables are independent, so a low p-value will reject independence and act as a measure of correlation. We also calculate Cramer's V as a scaled measure of association for nominal variables.

Table 4 shows the Cramer's V measure with the (p-value) below in each cell.

Table 4:

Outcome	Gender treatment	Scenario treatment	Recent Support Experience	Subject Age	Subject Gender
Overall Customer Support	0.01 (0.87)	0.38 (0.00)	0.18 (0.00)	0.09 (0.16)	0 (1)
Support Rep Understanding	0.00 (1)	0.45 (0.00)	0.03 (0.59)	0.16 (0.01)	0 (1)
Support Rep Knowledge	0.05 (0.40)	0.44 (0.00)	0.18 (0.00)	0.15 (0.02)	0.05 (0.39)
Request Support Rep Again	0.00 (1)	0.59 (0.00)	0.12 (0.06)	0.08 (0.22)	0 (1)
Gift Card	0.04 (0.52)	0.46 (0.00)	0.04 (0.56)	0.2 (0.00)	0.02 (0.68)

Key findings we can extract from this correlation table include:

- While the scenario treatment indicating good or bad support is correlated with all outcome variables, the support representative gender treatment is not correlated with any outcomes.
- The pretreatment blocking variable indicating recent support experience is correlated with overall customer support and the rating for support representative knowledge level.
- Subject age appears to be correlated with the gift card outcome and perception of the support representative's understanding of the problem.
- The scenario treatment is correlated with outcomes. Test subjects noticed the quality of support changed and adjusted ratings accordingly.

Checking Covariates for Balance

We verified the *ceteris paribus* assumption, that all other things are equal except for the treatment effect. For each of our treatment variables we tested the null model of the treatment outcome against a model which included the treatment effect and covariates. The additional model features did not increase our ability to predict whether someone was in the treatment or control group for the treatment. We were satisfied with the randomness of our experiment data.

```
Analysis of Variance Table

Model 1: D_Female ~ 1 + Recent_Support_Experience_Pos + Subject_Age_Block +
    Subject_Gender_Female
Model 2: D_Female ~ 1
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1    262 65.944
2    265 66.440 -3  -0.49625 0.6572 0.5791
```

```
Analysis of Variance Table

Model 1: D_Bad_Support_Scenario ~ 1 + Recent_Support_Experience_Pos +
    Subject_Age_Block + Subject_Gender_Female
Model 2: D_Bad_Support_Scenario ~ 1
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1    262 66.244
2    265 66.316 -3  -0.07156 0.0943 0.9631
```

Attention Check

Treatment was applied to the survey respondents via the support scenario transcript which included a male or female support representative and either a good or bad support scenario. To measure whether respondents paid attention to the treatment, we included a question to verify that the subjects had read the scenarios. A small number of respondents failed this test.

Table 5:

Fail	Pass
28	238

We also validated our attention check measure using survey timing information provided by Lucid. For each respondent, Lucid provided the time in seconds to complete the survey. We performed an independent samples t-test to test the hypothesis that the mean completion time is the same for subjects who paid attention to the transcript and those who did not. The test rejected the hypothesis, indicating that there is a statistically significant difference in completion times for subjects who passed our attention check and those who failed.

```
        Welch Two Sample t-test

data:  Duration..Seconds. by Compliance_Response_Ind
t = -2.5645, df = 97.344, p-value = 0.01186
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -120.24044  -15.32678
sample estimates:
mean in group 0 mean in group 1
       83.03571       150.81933
```

The number of subjects who possibly may not have paid attention to treatment is small compared with those who passed. It is also possible that although they failed to pay close attention to the scenario, the respondents may still have noticed the gender of the support representative.

Hypothesis Testing for Outcome Variables

The models testing our hypotheses were applied to each of the outcome variables.

Notation

F_i : gender treatment, 1 for female support representative, 0 for male
B_i : scenario treatment, 1 for bad support, 0 for good support
P_i : pretreatment variable measuring predisposition to support
A_i : subject age factor
G_i : subject gender

In each case we calculated robust standard errors to estimate our confidence intervals.

We used the binary outcome variables in the analysis below. We also performed hypothesis analysis using the numeric amount of the gift card, but the results were very similar to the binary indicator, since our respondents only used three of the five possible gift amounts.

Gender Based Ratings Per Scenario

We tested the hypothesis that female support representatives receive worse ratings than male representatives. Holding constant the value for scenario allowed us to model results for the gender treatment for each outcome variable.

$$Y_{i,B=1} = aF_i + u_i$$

Table 6 shows the results of our first hypothesis for each outcome.

Table 6:

Outcome	Variable	Bad Support ATE	Confidence Interval	Good Support ATE	Confidence Interval
Overall Customer	Gender effect	-0.0025	(-0.1812, 0.1762)	0.0351	(-0.0732, 0.1434)

Support					
Support Rep Understanding	Gender effect	0.0078	(-0.1656, 0.1812)	8e-04	(-0.0395, 0.0411)
Support Rep Knowledge	Gender effect	0.0404	(-0.132, 0.2127)	-0.1462	(-0.2767, -0.0157)
Request Support Rep Again	Gender effect	-0.0018	(-0.1624, 0.1588)	-0.0065	(-0.1228, 0.1097)
Gift Card	Gender effect	0.0875	(-0.0849, 0.2599)	0.0098	(-0.1182, 0.1378)

We draw several conclusions from this data:

- Female support representatives received slightly worse overall customer satisfaction ratings for poor support, and slightly better ratings for good support.
- Female support representatives were slightly more likely than males to receive a gift card regardless of the quality of support.
- The ability of female support representatives to understand the technical problem was rated slightly higher for both good and bad support.
- But it was less likely that our survey subjects would request the female support representative to help them again.

The results just described were not statistically significant based on the confidence intervals. Therefore, we cannot reject the null hypothesis that support representative gender does not impact customer support ratings for these outcomes.

The data also shows that the knowledge of female support reps providing good customer support is valued worse than their male counterparts providing the same level of support.

This value was statistically significant. Given that this is the only significant result in this table we are mindful of the possibility that it may still be an effect due to random chance.

Treatment by Treatment Interaction

The interaction between scenario type and gender treatments generated by modeling treatment by treatment interactions showed how much more gender mattered for bad or good support.

$$Y_i = a + bB_i + cF_i + d(B_iF_i) + u_i$$

Table 7 shows the results of our second hypothesis for each outcome.

Table 7:

Outcome	**Variable**	**Interaction**	**Confidence Interval**

Overall Customer Support	Rep Gender x Support Scenario	-0.0377	(-0.2466, 0.1713)
Support Rep Understanding	Rep Gender x Support Scenario	0.007	(-0.171, 0.185)
Support Rep Knowledge	Rep Gender x Support Scenario	0.1866	(-0.0296, 0.4028)
Request Support Rep Again	Rep Gender x Support Scenario	0.0048	(-0.1935, 0.203)
Gift Card	Rep Gender x Support Scenario	0.0777	(-0.137, 0.2925)

Although the results were not statistically significant, we draw several conclusions from this data:

- Overall customer support rating declines for female support representatives as support quality becomes poor.
- The other outcomes all had positive interaction coefficients. For instance, support representative knowledge rating improves for female support reps as support quality becomes poor.
- A small improvement such as the interaction effect for Request Support Rep Again would not be large enough to outweigh the negative gender effect.

Subject Gender and Age Covariates

In an effort to improve the robust standard errors of our estimates we tested including the age and gender of our test subjects as covariates. Precision of our estimates was not improved. We therefore estimated treatment effects without these covariates.

Heterogeneous Treatment Effects

We evaluated the hypothesis that gender treatment may have different effects based on the gender and age of the subject. Through regression analysis, we used interaction terms between the treatment effect and covariates.

Subject Gender Analysis

$$Y_{i,B=1} = a + bF_i + cG_i + d(F_iG_i) + u_i$$

Table 8 shows the results of our third hypothesis for each outcome.

Table 8:

Outcome	Variable	Bad Support HTE	Confidence Interval	Good Support HTE	Confidence Interval
Overall Customer Support	Agent Gender x Subject Gender	0.1179	(-0.2445, 0.4802)	-0.0686	(-0.2842, 0.147)

Support Rep Understanding	Agent Gender x Subject Gender	0.2699	(-0.0881, 0.6279)	0.0017	(-0.0808, 0.0842)
Support Rep Knowledge	Agent Gender x Subject Gender	0.1363	(-0.2034, 0.4759)	0.1053	(-0.1504, 0.361)
Request Support Rep Again	Agent Gender x Subject Gender	0.1646	(-0.1603, 0.4895	-0.0595	(-0.2949, 0.176)
Gift Card	Agent Gender x Subject Gender	0.2615	(-0.0969, 0.6199)	-0.0444	(-0.3046, 0.2159)

While we did not find a statistically significant heterogeneous treatment between genders, we found it interesting that all bad support scenario interaction coefficients are positive. There may be some indication here that female subjects are more likely to respond positively to a female agent, in the event that the agent's performance is poor.

Subject Age Analysis

$$Y_{i,B=1} = a + bF_i + cA_i + d(F_iA_i) + u_i$$

Table 9 below shows the results of our fourth hypothesis for each outcome.

Table 9:

Outcome	**Variable**	**Bad Support HTE**	**Confidence Interval**	**Good Support HTE**	**Confidence Interval**
Overall Customer Support	Agent Gender x Age Group	-0.1608	(-0.5154, 0.1938)	0.0077	(-0.2127, 0.2281)
Support Rep Understanding	Agent Gender x Age Group	-0.0167	(-0.3568, 0.3234)	-0.0025	(-0.0758, 0.0708)
Support Rep Knowledge	Agent Gender x Age Group	-0.2638	(-0.5986, 0.0711)	-0.0571	(-0.3172, 0.203)
Request Support Rep Again	Agent Gender x Age Group	-0.0126	(-0.3361, 0.3109)	0.1545	(-0.0815, 0.3905)
Gift Card	Agent Gender x Age Group	-0.0485	(-0.3875, 0.2904)	0.0621	(-0.1933, 0.3176)

While we did not find a statistically significant heterogeneous treatment between ages, we found interestingly that all bad support scenario interaction coefficients are negative.

There may be some indication here that older subjects are more likely to respond negatively to a female agent, in the event that the agent's performance is poor.

Single Factor Analysis

Taking our analysis of the gender treatment effect a step further we decided to collapse the scenario treatment factor to average over the good and bad support scenarios. We then retested the hypothesis that female support representatives receive worse ratings than male.

We used the following models to test the gender treatment alone, then combined gender with a dummy variable for the scenario treatment, and finally added the pretreatment blocking question indicating predisposition to customer support.

$$Y_i = aF_i + u_i$$
$$Y_i = aF_i + bB_i + u_i$$
$$Y_i = aF_i + bB_i + cP_i + u_i$$

Table 10 shows the treatment effect coefficient and robust standard error for each outcome using the single factor models just described.

Table 10:

	Gender treatment	Gender + scenario	Gender + scenario + pretreatment
Overall Customer Support	0.017 (0.055)	0.017 (0.051)	0.027 (0.050)
Support Rep Understanding	0.004 (0.048)	0.004 (0.043)	0.006 (0.043)
Support Rep Knowledge	-0.059 (0.060)	-0.058 (0.054)	-0.047 (0.053)
Request Support Rep Again	-0.005 (0.061)	-0.004 (0.049)	0.003 (0.049)
Gift Card	0.046 (0.060)	0.047 (0.053)	0.050 (0.053)

From these results we conclude that:

- The knowledge of the female support representative is rated less highly than that of male support representatives.
- Female support representatives are more likely to be offered a monetary bonus by customers than male.

These results are not statistically significant. We also see very little improvement in the standard error from adding the covariates.

Ordinal Logistic Regression

Using the same approach to collapse the treatment to a single gender factor, we re-ran the regression using ordinal logistic regression on the full Likert scale outcome values rather than the binarized form of the outcomes. We used the same models as the previous linear regression test, starting with gender treatment, then adding the additional covariates for scenario and pretreatment variable.

$$Y_i = aF_i + u_i$$
$$Y_i = aF_i + bB_i + u_i$$
$$Y_i = aF_i + bB_i + cP_i + u_i$$

Table 11 shows the gender treatment coefficient and standard error for each outcome using these models.

Table 11:

	Gender treatment	Gender + scenario	Gender + scenario + pretreatment
Overall Customer Support	0.096 (0.230)	0.139 (0.246)	0.171 (0.248)
Support Rep Understanding	-0.124 (0.223)	-0.162 (0.236)	-0.139 (0.237)
Support Rep Knowledge	-0.363 (0.223)	-0.434 (0.229)	-0.401 (0.231)
Request Support Rep Again	-0.008 (0.241)	0.043 (0.275)	0.085 (0.277)
Gift Card	-0.099 (0.222)	-0.106 (0.230)	-0.097 (0.231)

The pretreatment covariate does not appear to improve the standard error of the treatment estimate.

The two highlighted results were both statistically significant with a negative coefficient for the support representative knowledge rating.

These results compared to the previous linear regression table suggest that differences in ratings may be too subtle to capture with binarized outcomes. A small difference between very positive and positive would not be reflected in results which group those two values into the same binary outcome.

Randomization Inference

We performed randomization inference on the survey results to estimate whether individual treatment effects could occur by random chance. For each of our outcome variables, our randomization inference analysis tests the sharp null hypothesis, which is that the gender treatment has no effect on the measured outcome of interest for any of our subjects.

For each outcome variable, we performed randomization inference on the subset of bad support data, the subset of good support data, and the combined aggregate data. For the separate subsets of bad support or good support data, we simulated blocked randomization, where blocking is by the pre-treatment question, which asked the subject to rate their most recent technical support experience. Subjects who responded positively to this pre-treatment question were placed with the same block, and subjects who responded neutral or negatively were placed in a different block. Simulated random assignment was then performed within each block, where subjects were repeatedly re-assigned to control (male gender support rep) and treatment (female gender support rep) 100,000 times.

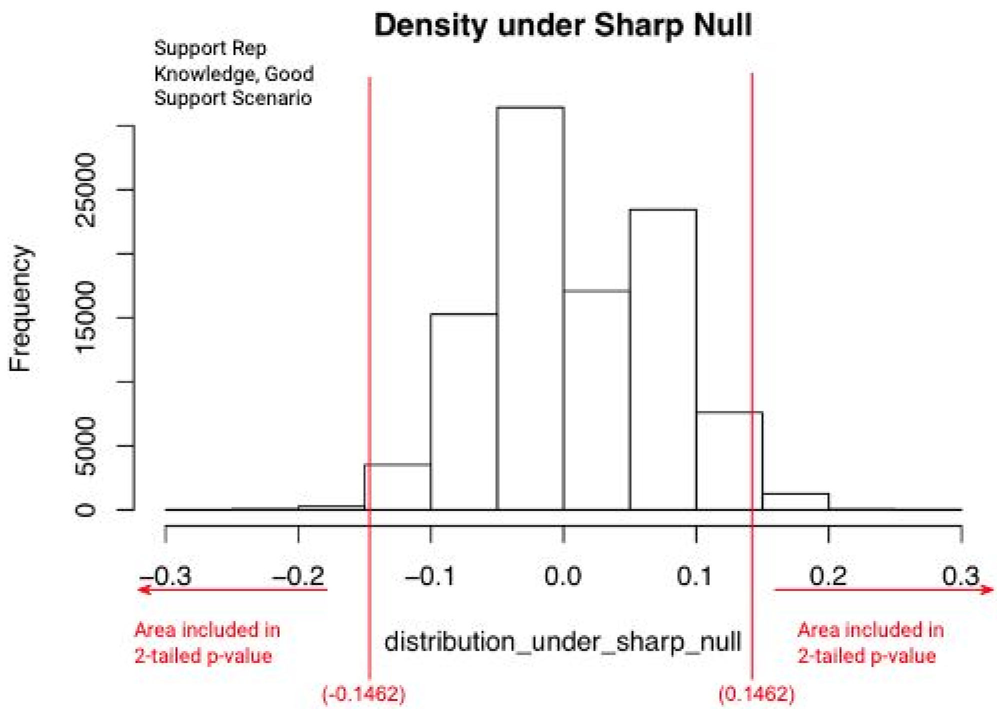
For the combined aggregate data, blocked randomization was simulated for nested blocks. Subjects are first separated into blocks by the pre-treatment question responses. Within each pre-treatment question response block, subjects are then blocked again by assignment to good support scenario and bad support scenario. Random assignment to the gender treatment was then simulated 100,000 times within each pretreatment question by support scenario inner block.

To maintain the accuracy of the randomization inference analysis, the data was sorted so that subjects of the same block assignment were grouped together. Repeated random assignment to the gender treatment was then simulated within blocks, where each iteration of the random assignment resulted in control and treatment counts that match those in Table 2, which display the actual distribution of counts of subject in the experiment assigned to each block and support scenario.

Within each iteration of random assignment, simulated average treatment effects were calculated using the random assignment result and the actual observed data for the outcome variable. Under the sharp null hypothesis, we expect the treatment to have no individual effect, which means that all observed outcomes represent the potential outcome of every subject in both treatment and control conditions and the simulated ATEs are thus due to random chance. Two-tailed p-values were then calculated by finding the proportion of simulated average treatment effects out of 100,000 that were as extreme as the absolute value of the estimated treatment effect from the actual assignment to treatment or control observed in the experimental data.

Figure 7 shows one example of the distribution of simulated ATEs for the outcome variable, Support Representative's Knowledge Rating, in the bad support scenario only. For this outcome variable, the estimated average treatment effect from the data is -0.1462 and the proportion of simulated ATEs due to random chance that are as extreme as the measured ATE is 0.0258, which represents the two-tailed p-value.

Figure 7:



Partitioning by Good and Bad Support Scenario

Table 12:

Outcome	Bad Support ATE	2-tailed p-value	Good Support ATE	2-tailed p-value
Overall Customer Support	-0.0025	~ 1	0.0351	0.6001
Support Rep Understanding	0.0078	~ 1	8e-04	~ 1
Support Rep Knowledge	0.0404	0.7153	-0.1462	0.0258
Request Support Rep Again	-0.0018	~ 1	-0.0065	~1
Gift Card	0.0875	0.3524	0.0098	~1

First we performed randomization inference on the good support scenario data and the bad support scenario data separately for each of the 5 outcome variables. Each of the 10 instances of randomization inference consisted of repeated block randomization, with pre-treatment question responses as the blocking variable. Based on the summary of results in Table 12, only one of the 10 instances of randomization inference analysis produced a significant result at the 0.05 significance level. The variable, Support Rep Knowledge in the good support scenario, is the same variable that exhibited a statistically significant result in both linear regression and logistic regression analysis.

Collapsing Good/Bad Support Scenario Data

Next, we performed randomization inference on the entire combined data set. The simulated randomization still takes good or bad support scenario into consideration, but does so as nested blocks within the pretreatment question blocking variable. In the results in Table 13, we see that none of the outcome variables result in a significant treatment effect since all p-value from simulated treatment effects under the sharp null hypothesis were insufficiently small enough

Table 13:

Outcome	Combined ATE	2-tailed p-value
Overall Customer Support	0.0167	0.7587
Support Rep Understanding	0.0036	~ 1
Support Rep Knowledge	-0.0586	0.3269
Request Support Rep Again	-0.0052	~ 1
Gift Card	0.0459	0.3886

Based on all 15 instances of randomization inference, 14 of which resulted in statistically insignificant p-values, we conclude that we fail to reject the sharp null hypothesis, which states that the gender treatment had no individual effect on the survey ratings of any of our subjects. For many of the outcome variables, the two-tailed p-values were rather large, with half of them close to 1, which signifies that random chance could lead to observed outcomes that could result in ATEs greater than those we observed.

Multiple Comparisons

As we designed our experiment we were cognizant of the consequences for the significance of our results in utilizing nested blocking and multiple outcomes.

Our gender and age analysis was performed without nesting the blocks to minimize the number of comparisons being performed and avoid a fishing expedition which would require Bonferroni correction of our p-value estimates.

Principal Component Analysis of Outcome Variables

Dimensionality reduction of the outcome variables was attempted to mitigate the risk that there may be similarities between the outcomes being captured. Initially, we reduced all five outcomes to a single component. We then used analysis of the correlation between outcomes to select the three outcomes with the highest correlation to reduce down to a single component.

We modeled the new outcome variable using gender treatment and the pretreatment covariate.

$$Y_i = aF_i + bP_i + u_i$$

Tables 14 and 15 below provide a comparison of the gender treatment effect for the most highly correlated outcomes and the effect after dimensionality reduction. We ran two tests of this approach first using data from the bad support scenario, then using the combined set of data from both good and bad support.

Table 14:

Only bad support data	Treatment effect coefficients and robust standard errors				
	Overall Customer Support	Support Rep Knowledge	Request Support Rep Again	PCA on 3 outcomes	PCA on 5 outcomes
Gender effect	0.028 (0.091)	0.075 (0.087)	0.031 (0.082)	-0.079 (0.127)	-0.115 (0.154)

Table 15:

Averaged good/bad support data	Treatment effect coefficients and robust standard errors				
	Overall Customer Support	Support Rep Knowledge	Request Support Rep Again	PCA on 3 outcomes	PCA on 5 outcomes
Gender effect	0.026 (0.055)	-0.049 (0.060)	0.001 (0.061)	0.015 (0.088)	-0.010 (0.102)

We found that models based on principal components did not substantially improve the magnitude of the estimates or reduce the standard error compared with models based on the original outcomes. The correlation analysis suggests that while there is some correlation between three of the outcome variables, there is too much variance in the outcomes projected onto a lower dimensional plane to provide better estimates.

Feynman Analysis

"if you're doing an experiment, you should report everything that you think might make it invalid - not only what you think is right about it" - Richard Feynman.

Our experiment was deliberately implemented online to allow the gender of the support representative to be varied without introducing other elements of variation that would have created noise if live actors were involved, such as race and age. This design decision means that our subjects were sourced from people who respond to online surveys. How well these people represent the population at large impacts the generalizability of our results. This is mitigated to some extent by the fact that we are researching effects related to technical support and benefit from using respondents familiar with online technology. However, there are still many people familiar with online technology and receiving technical support who are not active participants in online surveys.

Whether or not our gender treatment was noticed by survey respondents is another potential concern. The changes we made following the pilot study should increase the likelihood of gender treatment being noticed, but it is difficult to know for sure in our experiment.

Perhaps it is not awareness of gender differences that triggers bias, but voice or other gender-related communication characteristics that induce bias. This would be difficult to capture in an online experiment, but would make for interesting future research using another methodology.

Binarization of our Likert scale survey responses may have obscured some finer grained variation in responses. If our subjects provided a slightly below positive rating for females than males on some survey questions, which our histograms and ordinal logistic regression results suggest is possible, then we may not have captured that variation in our binary analysis.

It is also possible that the changes we made in our experimental design following the pilot study presented more difficulty in detecting statistically significant results in our actual experiment. Our 2x2 factorial design, aimed to study the effect of gender in both good support and bad support scenarios, was highly motivated by our speculation that the gender treatment effect would be greater in the bad support scenario case. This was also the reason that our pilot focused primarily on the gender treatment in the bad support scenario in order to assess whether or not we could detect a significant treatment effect in the later experiment. However, the expectations from the analysis of our pilot study data did not match the results of our experiment, neither in the size nor sign of the ATEs or number of outcome variables that would exhibit significant ATEs. The ATEs we found in the pilot study from 28 subjects were all larger in magnitude than those found from the survey data collected from 126 subjects assigned to bad support scenario. Perhaps the change we made to the bad support scenario script following the pilot, where the support representative ends the chat by offering to open an escalation ticket, is responsible for the decrease is ATE sizes in the experiment. In the original bad support scenario, the agent ends the chat without being able to find a proper resolution, thereby appearing far less competent than the agent who opens the escalation ticket in the revised version. It is possible that in order to see a larger gender treatment effect, we should have designed a bad support scenario script where the the agent performance was even worse than that portrayed in the pilot script.

Looking Ahead

Many organizations have an interest in both ensuring accurate results from customer satisfaction surveys and encouraging workplace diversity. Providing a common repository for organizations to share their results would allow data to be pooled to generate better estimates. Pooling across multiple studies to create a combined estimate allows us to incorporate information from a range of different studies with significant or insignificant treatment effects. We can combine data available from multiple studies into a single estimate by taking the estimates of the different studies and calculating an average weighted by the inverse of the standard error squared. Such an initiative could be promoted through programs and organizations encouraging Women in Technology.

Conclusion

If gender of support reps does have an effect on the reported level of satisfaction by customers, we were unable to observe definitive evidence of it in our experiment. Some of our results do suggest though that there may be an effect of gender on the perceived knowledge of female technical support representatives, particularly in cases where support provided is less than perfect. This effect may be too subtle to be reflected in binary outcome ratings and only appears when ratings allow subjects to give more varied responses.

This project was a great learning experience for our team. Our team slack channel is full of discussions on such topics as multiple comparisons, bias, factorial and blocking design, and coefficient interpretation. We look forward to applying these skills in future work.

References

1. Gallup, “Business Benefits of Gender Diversity”, url: https://www.gallup.com/workplace/236543/business-benefits-gender-diversity.aspx

2. McKinsey, “Women in the Workplace - 2018”, url: https://www.mckinsey.com/featured-insights/gender-equality/women-in-the-workplace-2018

3. Northwestern Institute for Policy Research, “Why Do So Few Women Hold Positions of Power”, url: https://www.ipr.northwestern.edu/about/news/2016/why-so-few-women-hold-positions-of-power. html

4. Dalian F Flake, “When Should Employers Be Liable for Factoring Biased Customer Feedback into Employment Decisions?”, Minnesota Law Review. May 2018, Vol. 102 Issue 5

5. Kimberly McGee, “The influence of gender, and race/ethnicity on advancement in information technology (IT)”, Information and Organization. March 2018. 28(1)

6. David R Hekman, Karl Aquino, Bradley P. Owens, Terence R. Mitchell, Pauline Schilpzand, Keith Leavitt, "An Examination of Whether and How Racial And Gender Biases Influence Customer Satisfaction", Academy of Management Journal. 2010. Vol 53, No 2., url: https://journals.aom.org/doi/10.5465/amj.2010.49388763

Appendix

Male, Poor Support

👦**Henry**: Good Evening! My name is Henry. What can I help you with today?

Customer: I am having trouble connecting my laptop to the WiFi, just like my phone.

👦**Henry**: I am sorry you are experiencing this issue. One moment please... I am going to remotely check the WiFi connection.

(2 min later)

👦**Henry**: My diagnostic test did not show any WiFi connection issues. Everything appears to be working well.

Customer: I am talking about a problem with the laptop connecting to WiFi.

👦**Henry**: I am not sure what you mean. The diagnostic test did not show any issues. Are you having a problem with the WiFi connection?

Customer: I know the WiFi connection is fine because my phone is connected to it and works fine. The issue is that my laptop won't connect to the WiFi.

👦**Henry**: OK, I get it. It's probably a network configuration issue on the laptop. What is the laptop's operating system?

Customer: Linux

👦**Henry**: Let me check if we have any documentation for troubleshooting Linux WiFi.

(2 min later)

Customer: Are you still there?

👦**Henry**: Yes I am. We only have steps for Windows and Mac. I will need to escalate this to Level 2 support, and they will contact you.

Customer: How long will this take?

👦**Henry**: I'm not sure exactly.

Customer: I have a demo tomorrow morning and I really need the WiFi working so that I can use my laptop to present.

👦**Henry**: I will note that on the ticket so Level 2 is aware of the situation.

Customer: OK

👦**Henry**: Is there anything else I can help you with today?

Customer: No, that is all.

🧑**Henry:** It has been a pleasure working with you today. Please do not hesitate to reach out on our live chat should you have any further issues. Thank you and have a wonderful day!

Female, Poor Support

👩**Angela:** Good Evening! My name is Angela. What can I help you with today?

Customer: I am having trouble connecting my laptop to the WiFi, just like my phone.

👩**Angela:** I am sorry you are experiencing this issue. One moment please... I am going to remotely check the WiFi connection.

(2 min later)

👩**Angela:** The diagnostic test did not show any WiFi connection issues. Everything appears to be working well.

Customer: I am talking about a problem with the laptop connecting to WiFi.

👩**Angela:** I am not sure what you mean. The diagnostic test did not show any issues. Are you having a problem with the WiFi connection?

Customer: I know the WiFi connection is fine because my phone is connected to it and works fine. The issue is that my laptop won't connect to the wifi.

👩**Angela:** OK, I get it. It's probably a network configuration issue on the laptop. What is the laptop's operating system?

Customer: Linux

👩**Angela:** Let me check if we have any documentation for troubleshooting Linux WiFi.

(2 min later)

Customer: Are you still there?

👩**Angela:** Yes I am. We only have steps for Windows and Mac. I will need to escalate this to Level 2 support, and they will contact you.

Customer: How long will this take?

👩**Angela:** I'm not sure exactly.

Customer: I have a demo tomorrow morning and I really need the WiFi working so that I can use my laptop to present.

👩**Angela:** I will note that on the ticket so Level 2 is aware of the situation.

Customer: OK

👩**Angela:** Is there anything else I can help you with today?

Customer: No, that is all.

👩**Angela:** It has been a pleasure working with you today. Please do not hesitate to reach out on our live chat should you have any further issues. Thank you and have a wonderful day!

Male, Good Support

🧑**Henry:** Good Evening! My name is Henry. What can I help you with today?

Customer: I am having trouble connecting my laptop to the WiFi, just like my phone.

🧑**Henry:** I am sorry you are experiencing this issue. If I understand correctly, your phone is using the WiFi connection without issues, but your laptop will not connect. Can you confirm?

Customer: Yes, my phone WiFi is working but my laptop will not connect.

🧑**Henry:** OK, I get it. This is probably a network configuration issue on the laptop. What is the laptop's operating system?

Customer: Linux

🧑**Henry:** I see. We have steps for Windows and Mac but not Linux. Let me put some steps together for you.

Customer: How long will this take?

🧑**Henry:** Not long, I am very familiar with Linux. Just need to write up the steps and attach screenshots. Do you mind being on hold for a couple of minutes while I do that?

Customer: Yes, I will be here.

(2 min later)

🧑**Henry:** Hello - are you still there? I am attaching the troubleshooting document. I will wait on the line to confirm the issue has been fixed. 📎`attachment`

Customer: Thanks! Taking a look now.

(3 min later)

Customer: It's working! My laptop is now on WiFi.

🧑**Henry:** Glad to hear. Is there anything else I can help you with today?

Customer: No, thank you for your help.

🧒**Henry:** It has been a pleasure working with you today. Please do not hesitate to reach out on our live chat should you have any further issues. Thank you and have a wonderful day!

Female, Good Support

👩**Angela:** Good Evening! My name is Angela. What can I help you with today?

Customer: I am having trouble connecting my laptop to the WiFi, just like my phone.

👩**Angela:** I am sorry you are experiencing this issue. If I understand correctly, your phone is using the WiFi connection without issues, but your laptop will not connect. Can you confirm?

Customer: Yes, my phone WiFi is working but my laptop will not connect.

👩**Angela:** OK, I get it. This is probably a network configuration issue on the laptop. What is the laptop's operating system?

Customer: Linux

👩**Angela:** I see. We have steps for Windows and Mac but not Linux. Let me put some steps together for you.

Customer: How long will this take?

👩**Angela:** Not long, I am very familiar with Linux. Just need to write up the steps and attach screenshots. Do you mind being on hold for a couple of minutes while I do that?

Customer: Yes, I will be here.

(2 min later)

👩**Angela:** Hello - are you still there? I am attaching the troubleshooting document. I will wait on the line to confirm the issue has been fixed. 📎`attachment`

Customer: Thanks! Taking a look now.

(3 min later)

Customer: It's working! My laptop is now on WiFi.

👩**Angela:** Glad to hear. Is there anything else I can help you with today?

Customer: No, thank you for your help.

👩**Angela:** It has been a pleasure working with you today. Please do not hesitate to reach out on our live chat should you have any further issues. Thank you and have a wonderful day!