

W241 Final Project: Perception of Tap Water and Consumption of Bottled Water

Andrea, Scott, David

8/14/2018

I. Objective and Overview

Clean tap water is ubiquitous and easily accessible by any person living in a prosperous country. However, the general population might not be as aware of its benefits and overall cleanliness- with most choosing to purchase and drink bottled water. Oftentimes, individuals carry certain perspectives on tap water, which could affect their drinking habits, as well as their overall confidence in the tap water in their area.

The team wanted to study people's understanding and perception of tap water further. We wanted to test if providing a population with some baseline education on the benefits of tap water could change the bottled water to tap water consumption ratio or their overall sentiment about tap water. Should we present people with supplementary information on tap water, would there be a significant change in their attitudes towards tap water?

Statement of the Team's Hypothesis

The null hypothesis is: Informing people of tap water quality will not influence public perception of tap water and bottled water consumption.

The alternative hypothesis is: Informing people of tap water quality will increase public perception of tap water and decrease bottled water consumption.

II. Experimental Design and Methodology

1. Run a pilot for ~10 people to see if the process of surveying, presenting treatment videos and placebo videos is straightforward enough. The team also wanted to see if there would be any feedback on the phrasing of questions to ensure utmost clarity.
2. Gather 30-35 individuals each using social networks, professional networks, friends and family.
3. Randomize and group these subjects into treatment, control and placebo group.
4. Conduct a pre-treatment and post treatment survey over the course of two weeks. The survey contains questions about their perception of tap water and their drinking habits.
5. Data Gathering, Cleaning and Analysis of Results.

We created the survey through Qualtrics, which is explained in more detail in the next sections. The team agreed to distribute the surveys to the participants via individual email links, which was a convenient function from Qualtrics. The email body requested the participants to answer the survey truthfully, not to share their experience with anyone, and assured the participants that their email addresses would not be used for any other manner except for this experiment. All three groups received the same exact email.

A. Pilot Study

A pilot study was performed to identify errors that would negatively impact the results of the study. We selected 10 participants through our social network, and asked them for feedback on the question clarity, experiment process, and overall satisfaction with participating. We were able to retrieve extremely beneficial

feedback which was implemented into the full-scale experiment. The pilot study participants were not invited to participate in the experiment.

B. Subject Selection Criteria

We performed a power analysis to determine the sample size required for this experiment. Based on a power level of 0.8, a cohen's d value of 0.5 to indicate medium effect, and a significance level of 95%, we determined that each group would require approximately 64 participants. Unfortunately, due to the limitations of our social network and the class timeline, we were only able to recruit 35-36 participants for each group. This would lead to an underpowered experiment.

```
##
##      Two-sample t test power calculation
##
##              n = 63.76561
##              d = 0.5
##      sig.level = 0.05
##              power = 0.8
##      alternative = two.sided
##
## NOTE: n is number in *each* group
```

The subjects were selected from our personal network (colleagues, friends, etc.) Each author was tasked with recruiting 35 participants, we were able to recruit 107 participants. The participants were then assigned to treatment, control, and placebo groups. The randomization procedure is described in following sections.

C. Pre-treatment and Post-treatment

The Pre-treatment and Post-treatment surveys were practically identical, with the exception of the first two questions (Gender and Age) excluded in the Post-survey, and the potential treatment or placebo videos excluded in the Post-treatment survey. The surveys were sent five days apart to the participants, allowing sufficient time (under class constraints) for the treatment effects to develop. The participants were given three days to complete each survey, they received friendly follow-up emails every day they did not complete the surveys.

D. A Note on Spillover Effects

One of the key concerns that the team had was susceptibility to spillover effects, that is subjects' opinions and conversations about the study and the questions causing a significant impact on the experiment that we are conducting. In an effort to prevent this from occurring within our experiment, the team made sure to do the following things:

The team made an effort to gather participants who were not family members or who could easily converse about the videos. The email copy that we crafted was neutral in tone to ensure that we did not reveal anything about the experiment. The email also explicitly requests that the topics and questions within the survey are not discussed with anyone in the next few weeks to prevent swaying the opinions of people they may know who might be part of the pool.

Now that we've sufficiently tackled a potential pitfall of this experiment, we can now move on to discussing the survey conducted, as well as the videos we had selected for the experiment.

E. Treatment, Control, and Placebo

The treatment group, control group, and placebo group were all asked to answer a short survey regarding tap water and their opinion on the matter. Each participant was emailed an individual link where once they clicked it, they were directed to the survey which contained the questions below.

- 1) What is your gender?
- 2) What is your age?
- 3) How much water do you drink per day?
- 4) How much bottled water do you consume per day?
- 5) How much do you enjoy drinking tap water?
- 6) Which has better quality (tap or bottled water)?
- 7) Are you confident in the quality of tap water?
- 8) Are you confident in the quality of bottled water?

Questions 3 and 4 were fill in the blank questions using 4 different units of measurement, gallons, cups, liters, and ounces. The participant was asked to answer in their preferred unit of measurement. Question 5 showed a picture of a emoji face, with the ability to scroll down to make the face sad, and the ability to scroll up to make the face happy. The different faces provided results numerically from 1-5, where 5 is if they really enjoyed drinking tap water, and 1 where they did not enjoy it at all. Question 6 was a multiple choice question, which allowed the participant to chose either tap water, bottled water, or same quality. Question 7 and 8 were displayed on a slide grading scale (A+ to F), with provided numerical results from 1 to 13, which 1 being F and 13 being A+.

The treatment and placebo groups were provided a video to watch at the end of the survey, the control group did not see any video.

F. Treatment and Control Video

We made sure to select short and digestible treatment and control videos to ensure that our subjects wouldn't lose interest. The treatment video(https://www.youtube.com/watch?v=8MGsb7SrW_s&feature=youtu.be) was a 4-minute newsclip which highlights the benefits of tap water. This video discussed in detail the quality, regulation, and treatment process tap water goes through before being served to the public. It also compares tap water and bottled water quality and concludes the difference to be negligible. The placebo video(https://www.youtube.com/watch?v=zw_isTcSKSM&feature=youtu.be) didn't really have to do anything with tap water, but was water related in some way. We chose this video because it was a UC Berkeley informational video about water efficiency in agriculture, we thought it would not have any treatment effects, and was brief enough to keep the viewers focus.

III. Random Assignments, Cleaning, Processing and Initial Analysis

The group started by randomizing the emails and contact information that the team compiled to ensure that there was no bias in how subjects were assigned into the CONTROL, TREATMENT AND PLACEBO groups. You can see the code below:

A. Randomize assignments

```
## [1] "treatment" "control" "placebo" "treatment" "control"
## [6] "placebo" "treatment" "treatment" "control" "placebo"
## [11] "placebo" "control" "control" "treatment" "placebo"
## [16] "treatment" "treatment" "placebo" "control" "treatment"
## [21] "control" "treatment" "control" "placebo" "control"
## [26] "placebo" "control" "treatment" "placebo" "treatment"
```

```
## [31] "placebo" "control" "treatment" "treatment" "treatment"
## [36] "placebo" "treatment" "placebo" "control" "control"
## [41] "treatment" "placebo" "control" "control" "treatment"
## [46] "treatment" "treatment" "treatment" "placebo" "placebo"
## [51] "treatment" "placebo" "control" "control" "control"
## [56] "placebo" "control" "placebo" "treatment" "placebo"
## [61] "placebo" "treatment" "control" "control" "placebo"
## [66] "placebo" "treatment" "treatment" "control" "treatment"
## [71] "control" "control" "control" "placebo" "control"
## [76] "placebo" "placebo" "placebo" "control" "placebo"
## [81] "treatment" "treatment" "placebo" "treatment" "control"
## [86] "control" "control" "treatment" "placebo" "control"
## [91] "control" "placebo" "placebo" "treatment" "treatment"
## [96] "treatment" "control" "placebo" "placebo" "treatment"
## [101] "treatment" "placebo" "control" "control" "placebo"
```

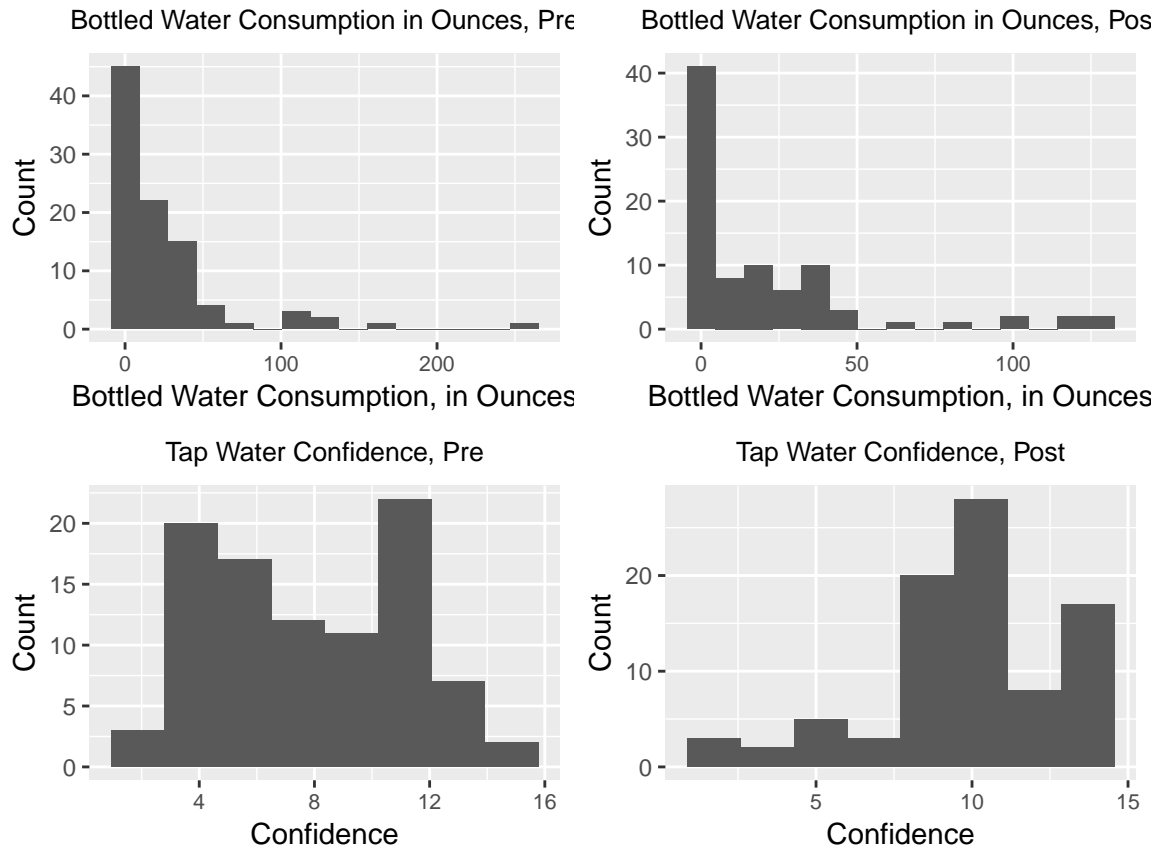
B. Cleaning

Next, once we had assigned the participants into groups and had run the experiment over a period of two weeks, we pulled the data from Qualtrics and did some data cleaning and manipulation that did the following:

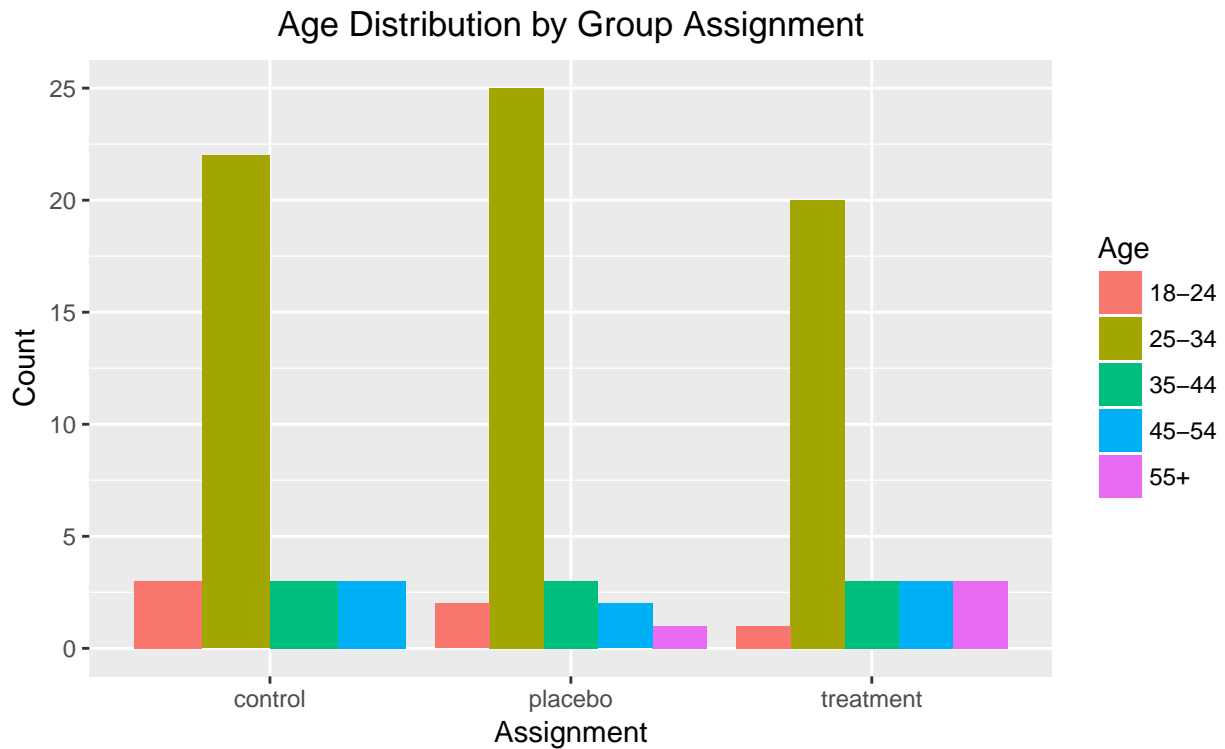
- 1) Removed NA's within the data set and changed them to 0 for easy calculation.
- 2) Merged csvs to make a master csv and therefore a cleaner dataframe to pull from.
- 3) Conversion of data point types for easy calculations.. 4) Some standard conversions of water measurements like ounces, gallons, liters and cups. We chose to standardize to ounces.

C. Visualizations

Finally, as another step of our cleaning and initial analysis, we wanted to run a few visualizations using histograms to further understand the responses of our survey-takers. You can see some of the examples below. The following graphs that we rendered focus on the answers gathered for questions around the consumption of bottled water and tap water confidence, pre and post treatment.

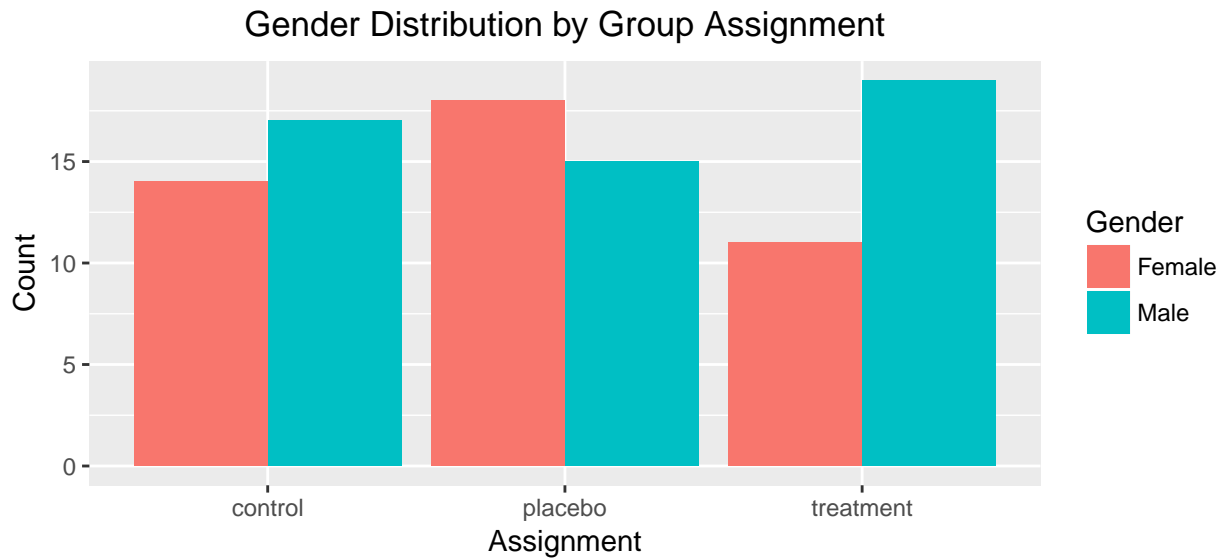


We also plotted the demographics of the sample population (Age and Gender).



The Age Distribution by Group Assignment figure shows that the most sampled age group is from 25 to 34.

This is due to the fact that the people we sampled are from people we know, who have similar age as us.



The Gender Distribution by Group Assignment figure shows the variation of gender for the treatment, control, and placebo groups. We didn't notice any one-sided heavy groups, therefore we proceeded with this variation.

D. Initial Analysis

After re-formatting the data, we will now perform initial analysis to observe potential significant results.

The metrics that the team wanted to focus on were the bottled water consumption metric, as well as the metrics that dealt with the individual's perception of tap water, like their confidence in it and their enjoyment.

E. CACE & ATE

For this experiment, we defined compliers as subjects that answered both the Pre-treatment survey and the Post-survey. There were 83 compliers out of the initial 94 people who responded to one of the surveys. Non-compliers were considered subjects who did not take one of the surveys. We excluded the subjects who did not complete either survey.

We then wanted to run a CACE or Average Treatment effect on the compliers. While at first glance there seemed to have been an positive effect on the consumption of bottled water, a simple t-test shows that the results weren't significant.

```
## [1] 4.721111
##
## Welch Two Sample t-test
##
## data: compliers$consumption_treatment_effect[compliers$assignment.x == and compliers$consumption_t.
## t = 0.72731, df = 47.413, p-value = 0.4706
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -8.334405 17.776627
## sample estimates:
## mean of x mean of y
## 4.6148148 -0.1062963
```

The team also looked at the Perception of Tap Water and the Enjoyment of it. The treatment effect on the perception of tap water is slightly negative, while the effect on the enjoyment of tap water is slightly positive. Both effects are statistically insignificant.

```
## [1] -0.6296296
##
## Welch Two Sample t-test
##
## data: compliers$confidence_treatment_effect[compliers$assignment.x ==  and compliers$confidence_treat
## t = -0.47869, df = 51.915, p-value = 0.6342
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -3.269128 2.009869
## sample estimates:
## mean of x mean of y
## 1.925926 2.555556
## [1] 0.1111111
##
## Welch Two Sample t-test
##
## data: compliers$enjoyment_treatment_effect[compliers$assignment.x ==  and compliers$enjoyment_treat
## t = 0.47803, df = 48.759, p-value = 0.6348
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.3560408 0.5782630
## sample estimates:
## mean of x mean of y
## 0.18518519 0.07407407
```

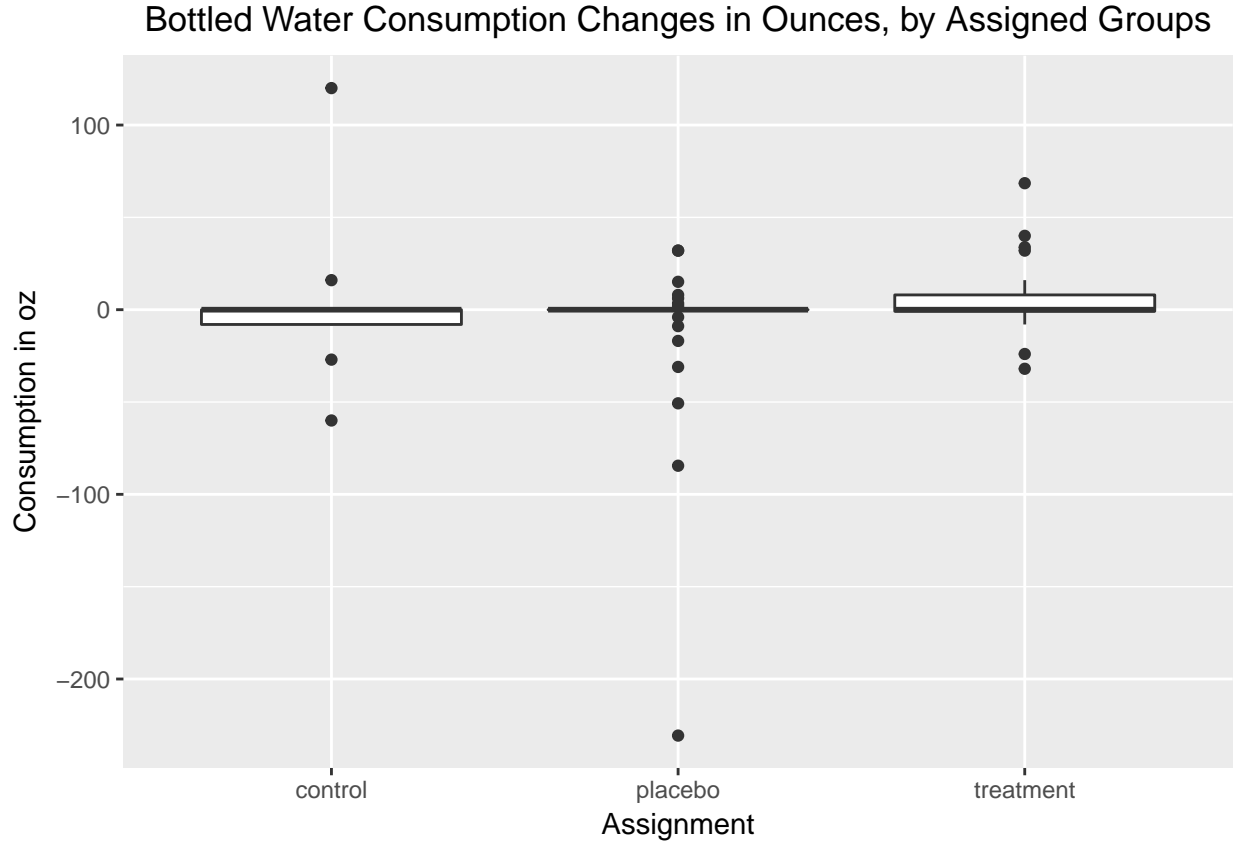
To take things further, the team ran a CACE on the Placebo and Treatment. This time, we took the question of whether or not they watched the accompanying video into account as another level of compliance.

```
## Treatment_Assignment Treated N Mean_bottled_consumption
## 1 Control No 27 -0.1062963
## 2 Treatment Yes 26 5.1000000
## 3 Treatment No 1 -8.0000000
## 4 Placebo Yes 26 -4.8200000
## 5 Placebo No 4 -67.5500000
```

What we saw here was a CACE Placebo of about -5.43 and a CACE of Treatment of 5.41. This is quite interesting because the placebo was meant to measure the potential outcomes of those who are untreated, and we expected this to have no impact (positive or negative) on the participants. Also, the treatment video increased bottled water consumption, which is the opposite of what we expected, since the video was selected to highlight the benefits of drinking tap water.

```
## [1] -5.438889
## [1] 5.406538
```

We thought the CACE of placebo being negative was unexpected, so we decided to plot the distributions of changes for all assigned groups. One thing that stands out is one particular extreme outlier. This outlier is sitting near -230; this, combined with our small sample sizes, means that our placebo's ITT down by a significant amount. Had we removed this outlier, the placebo would have had a much smaller CACE, close to 0.



IV. Modeling

A. Difference in Difference Treatment Effect

Because of our pre/post survey design, we can conduct Difference in Difference estimation with linear regression. This is useful because we waited a few days between the pre-treatment and post-treatment surveys, and using this technique allows us to mitigate the effects of extraneous factors that could have taken place during the 4 day waiting period.

i. Bottled Water Consumption

The following model shows the difference-in-difference on the Bottled Water Consumption of the treatment group and non-treatment groups (control, placebo). Treated is a dummy binary variable that indicates group assignment (1 for treatment, 0 for non-treatment). Post is a binary variable that indicates whether the measured consumption was pre-treatment, or post-treatment (1 for post, 0 for pre). We also included the interaction term Treated * post to measure the difference-in-difference.

$$\text{Consumption} = B_0 + B_1\text{Treated} + B_2\text{post} + B_3\text{Treated} * \text{post}$$

<i>Dependent variable:</i>			
Consumption of Bottled Water, in Ounces			
	(1)	(2)	(3)
Treated	-5.373 (4.929)	-5.357 (4.957)	-10.744 (6.964)
After		-2.325 (5.441)	-6.010 (7.403)
Treated x After			11.428 (9.912)
Constant	24.503*** (3.741)	25.588*** (5.036)	27.307*** (5.763)
Observations	177	177	177
R ²	0.005	0.006	0.011
Adjusted R ²	-0.001	-0.006	-0.006
Residual Std. Error	36.295 (df = 175)	36.381 (df = 174)	36.386 (df = 173)
F Statistic	0.847 (df = 1; 175)	0.512 (df = 2; 174)	0.658 (df = 3; 173)

Note:

*p<0.1; **p<0.05; ***p<0.01

The estimate coefficient for the interaction term is the DiD estimate. In this regression, the DiD for bottle water consumption pre and post treatment was $\beta_3 = 11.428$, and the result was not statistically significant.

ii. Tap Water Confidence

The following model shows the difference-in-difference on Tap Water Confidence of the treatment group and non-treatment groups (control, placebo). Once again, Treated is a dummy binary variable that indicates group assignment (1 for treatment, 0 for non-treatment). Post is a binary variable that indicates whether the measured consumption was pre-treatment, or post-treatment (1 for post, 0 for pre). We also included the interaction term Treated * post to measure the difference-in-difference.

$$\text{Confidence} = B_0 + B_1\text{TreatedDummy} + B_2\text{post} + B_3\text{TreatedDummy} * \text{post}$$

<i>Dependent variable:</i>			
Tap Water Confidence			
	(1)	(2)	(3)
Treated	0.881 (0.548)	0.866* (0.526)	0.992 (0.825)
After		2.110*** (0.479)	2.196*** (0.578)
Treated x After			-0.267 (1.045)
Constant	8.400*** (0.306)	7.415*** (0.386)	7.375*** (0.424)
Observations	177	177	177
R ²	0.015	0.113	0.113
Adjusted R ²	0.009	0.103	0.098
Residual Std. Error	3.355 (df = 175)	3.193 (df = 174)	3.202 (df = 173)
F Statistic	2.662 (df = 1; 175)	11.098*** (df = 2; 174)	7.381*** (df = 3; 173)

Note:

*p<0.1; **p<0.05; ***p<0.01

The estimate coefficient for the interaction term is the DiD estimate. In this regression, the DiD for Tap Water Confidence pre and post treatment was $\beta_3 = -0.267$, and the result was not statistically significant.

B. Heterogeneous Treatment Effect

We want to see whether the treatment effect on female participants is different from male participants. This heterogeneous treatment effect can be tested using linear regression.

i. Bottled Water Consumption

The following model shows whether heterogeneous treatment effect on male participants versus female participants, on the outcome variable Bottled Water Consumption. Treated is a dummy binary variable that indicates group assignment (1 for treatment, 0 for non-treatment). Female is a binary variable that indicates whether the participant is male or female (1 for female, 0 for male). We also included the interaction term Treated * Female to measure the heterogenous treatment effect.

$$\text{Consumption} = B_0 + B_1\text{Treated} + B_2\text{Female} + B_3\text{Treated} * \text{Female}$$

<i>Dependent variable:</i>			
Bottled Water Consumption			
	(1)	(2)	(3)
Treated	0.684 (7.053)	0.831 (7.430)	-6.979 (8.881)
Female		1.593 (7.502)	-4.086 (9.449)
Treated x Female			17.878 (15.518)
Constant	21.298*** (4.647)	20.501*** (6.719)	23.341*** (7.601)
Observations	83	83	83
R ²	0.0001	0.001	0.018
Adjusted R ²	-0.012	-0.024	-0.020
Residual Std. Error	32.270 (df = 81)	32.461 (df = 80)	32.389 (df = 79)
F Statistic	0.008 (df = 1; 81)	0.029 (df = 2; 80)	0.471 (df = 3; 79)

Note:

*p<0.1; **p<0.05; ***p<0.01

Based on the data presented, the HTE estimate is 17.878, and is not statistically significant.

ii. Tap Water Confidence

The following model shows whether heterogeneous treatment effect on male participants versus female participants, on tap water confidence. Once again, Treated is a dummy binary variable that indicates group assignment (1 for treatment, 0 for non-treatment). Female is a binary variable that indicates whether the participant is male or female (1 for female, 0 for male). We also included the interaction term Treated * Female to measure the heterogeneous treatment effect.

$$\text{Confidence} = B_0 + B_1\text{Treated} + B_2\text{Female} + B_3\text{Treated} * \text{Female}$$

<i>Dependent variable:</i>			
Tap Water Confidence			
	(1)	(2)	(3)
Treated	0.725 (0.641)	0.597 (0.628)	0.429 (0.736)
Female		-1.377** (0.623)	-1.500* (0.772)
Treated x Female			0.386 (1.342)
Constant	9.571*** (0.392)	10.260*** (0.407)	10.321*** (0.426)
Observations	83	83	83
R ²	0.015	0.075	0.076
Adjusted R ²	0.003	0.052	0.041
Residual Std. Error	2.809 (df = 81)	2.740 (df = 80)	2.755 (df = 79)
F Statistic	1.213 (df = 1; 81)	3.231** (df = 2; 80)	2.158* (df = 3; 79)

Note:

*p<0.1; **p<0.05; ***p<0.01

In this model, the HTE estimate is 0.386, and not statistically significant. We noticed that both the Female coefficient and the intercept are statistically significant. Keep in mind that the mean outcome of female participants in the control group is also $B_0 + B_2$, which means the female participants had a significant change in confidence in tap water.

C. Overall OLS Regression

i. Bottled Water Consumption

Past results indicated a lack of statistically significant results when looking at the Difference in Difference as well as the HTE. Out of curiosity, we want to account for every major variable in the OLS regression.

The following model predicts the bottled water consumption while accounting for the treatment, gender, tap water confidence before the treatment, and the age groups. Treated is a dummy binary variable that indicates group assignment (1 for treatment, 0 for non-treatment). Female is a binary variable that indicates whether the participant is male or female (1 for female, 0 for male). Pre Tap water Confidence is the participant's confidence in tap water measured in the pre-treatment survey. Age groups are a set of dummy variables that identifies the age bracket of the participant (only one of the dummy variables can be 1, the rest are 0)

$$\text{Consumption} = B_0 + B_1\text{Treated} + B_2\text{Female} + B_3\text{Pre Tap water Confidence} + B_4, 5, 6, 7\text{AgeGroups}$$

	<i>Dependent variable:</i>			
	Bottled Water Consumption			
	(1)	(2)	(3)	(4)
Treated	0.684 (7.053)	0.831 (7.430)	0.714 (8.162)	1.454 (8.192)
Female		1.593 (7.502)	1.483 (8.115)	-0.285 (8.377)
Tap water confidence, Pre			0.111 (1.185)	0.247 (1.253)
Age 18-24				33.455 (29.509)
Age 25-34				16.376** (8.059)
Age 35-44				4.424 (10.445)
Age 45-54				36.087** (17.763)
Constant	21.298*** (4.647)	20.501*** (6.719)	19.734** (9.147)	1.917 (13.604)
Observations	83	83	83	83
R ²	0.0001	0.001	0.001	0.073
Adjusted R ²	-0.012	-0.024	-0.037	-0.014
Residual Std. Error	32.270 (df = 81)	32.461 (df = 80)	32.663 (df = 79)	32.298 (df = 75)
F Statistic	0.008 (df = 1; 81)	0.029 (df = 2; 80)	0.023 (df = 3; 79)	0.839 (df = 7; 75)

Note:

*p<0.1; **p<0.05; ***p<0.01

In this model, the average treatment effect is 1.454. It is interesting to note that the highest consumption change is from the age group of 45 to 54, with a coefficient of 36.087. Only the age groups 25-34 and 45-54 are statistically significant.

ii. Tap Water Confidence

The following model predicts the tap water confidence while accounting for the treatment, gender, tap water confidence before the treatment, and the age groups. Treated is a dummy binary variable that indicates group assignment (1 for treatment, 0 for non-treatment). Female is a binary variable that indicates whether the participant is male or female (1 for female, 0 for male). Pre Tap water Confidence is the participant's confidence in tap water measured in the pre-treatment survey. Age groups are a set of dummy variables that identifies the age bracket of the participant (only one of the dummy variables can be 1, the rest are 0)

$$\text{Confidence} = B_0 + B_1\text{Treated} + B_2\text{Female} + B_3\text{Pre Bottled Consumption} + B_4, 5, 6, 7\text{Age Groups}$$

	<i>Dependent variable:</i>			
	Tap Water Confidence			
	(1)	(2)	(3)	(4)
Treated	0.725 (0.641)	0.597 (0.628)	0.518 (0.636)	0.461 (0.656)
Female		-1.377** (0.623)	-1.427** (0.621)	-1.382** (0.631)
Bottled Consumption, Pre			-0.008 (0.008)	-0.008 (0.008)
Age 18-24				-2.874 (2.778)
Age 25-34				-0.390 (1.621)
Age 35-44				-1.160 (1.935)
Age 45-54				0.163 (1.645)
Constant	9.571*** (0.392)	10.260*** (0.407)	10.493*** (0.471)	11.013*** (1.651)
Observations	83	83	83	83
R ²	0.015	0.075	0.087	0.135
Adjusted R ²	0.003	0.052	0.053	0.054
Residual Std. Error	2.809 (df = 81)	2.740 (df = 80)	2.738 (df = 79)	2.736 (df = 75)
F Statistic	1.213 (df = 1; 81)	3.231** (df = 2; 80)	2.518* (df = 3; 79)	1.670 (df = 7; 75)

Note:

*p<0.1; **p<0.05; ***p<0.01

The average treatment effect for this model is 0.461, and not statistically significant. Once again (similar to our HTE model), the female participants in the control group had a statistically significant change.

V. Conclusion

The team set out to determine the impacts of informing the public about the cleanliness and benefits of drinking tap water has on both the public's perception of tap water and their bottled water drinking habits. The experiment did not yield statistically significant results.

Ultimately, the team's findings on the impacts from the educational video were not statistically significant. The initial analysis seemed contradicting to our initial theory, with bottled water consumption increasing with treatment and the perception of tap water decreasing with treatment. This initial finding was discovered by calculation the CACE for both the treatment and placebo groups.

After further analysis, we concluded that the educational video did not have any statistically significant impact on either bottled water consumption or tap water perception. The analysis included running regression models to assess the heterogeneous treatment effect and the difference in difference effect since the study

took place over a period of time (5 days). The team did not find any statistically significant effects to tap water perception or bottled water consumption, therefore we cannot reject the null hypothesis.