

# “Don’t Procrastinate, Vaccinate”

An Experiment to Understand the Causal Effects of Raising Awareness on Preventive Healthcare

*W241: Experiments and Causality, Spring 2020: Alex Dessouky, Matt Kawa, Sonali Serro*  
*April 25, 2020*

## Abstract

In this study, we conduct an experiment to determine whether raising awareness around the risks of the seasonal flu causes more individuals to promote the prevention of infectious diseases. We apply two interventions, i) an influenza-awareness video, and ii) an influenza-awareness quiz. The experiment is conducted via the Qualtrics online survey platform and subjects are asked to allocate a \$1 donation between the National Foundation of Infectious Diseases (NFID) and a competing non-profit organization. Results of the analysis show a statistically significant conditional treatment effect of the influenza-awareness video on individuals with a higher propensity to take preventive health measures.

## 1 Introduction

Influenza spreads around the globe every year, causing about 3 to 5 million cases of severe illness and up to 650,000 deaths<sup>1</sup>. Approximately 20% of unvaccinated children and 10% of unvaccinated adults are infected each year<sup>2</sup>. Yearly vaccinations against influenza are recommended by the World Health Organization (WHO), especially for the high risk groups - the young, the old, pregnant women, and those with other health issues. However, according to the Centers for Disease Control and Prevention (CDC), fewer than half of Americans got the flu vaccine during the 2018-2019 flu season<sup>3</sup>.

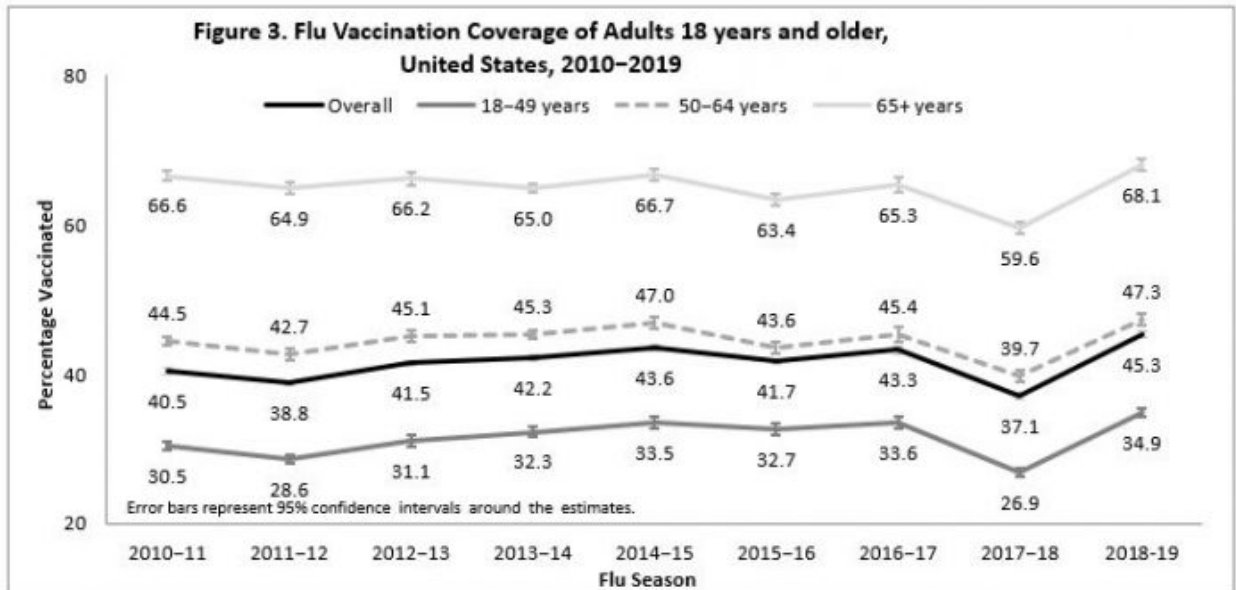


Fig. 1: Adult Flu Vaccination Coverage United States, 2010-2019, graph courtesy of the CDC

Recent studies<sup>4</sup> have shown that many Americans do not get the flu shot due to concerns around the efficacy of the vaccine and possible side effects, while others believe they are not likely to get the flu. Cost and accessibility do not appear to be a factor, given the growth in affordable flu clinics nationwide. Other studies indicate that younger individuals typically struggle with preventive healthcare, certain ethnic groups

(e.g. African Americans) are more susceptible to anti-vaccination rhetoric, and men are less likely than women to get vaccinated<sup>5</sup>.

## 2 Research Hypothesis

We hypothesize that lack of awareness is instrumental in creating this *vaccine lethargy* and hence the goal of the experiment was guided by the following research question: **Can raising awareness around the risks of influenza to oneself and others cause more individuals to promote effective prevention of infectious diseases?**

Our hypothesis is that an individual's propensity for preventive healthcare is an important covariate in predicting experimental outcomes, and that individuals with a higher propensity for preventive healthcare will be more responsive to the intervention.

### 2.1 Defining an Individual's Propensity for Preventive Healthcare

Our hypothesis of heterogeneous treatment effects was tested using relevant pre-treatment covariates that would systematically define subgroups of individuals responding differently to the intervention. We list the covariates below, and the rationale as to why they contribute to defining an individual's propensity for preventive healthcare.

- *Age*: For the 2018-19 season, flu vaccination coverage among adults between the ages of 18 and 49 years old was only 34.9%, as compared to 68.1% among adults older than 65<sup>3</sup>. One theory is that healthy young adults may be less-informed and perhaps do not prioritize preventive healthcare. This motivated the evaluation of treatment effect heterogeneity for adults <40 years old and adults ≥40 years old. Note that the mean age of the participants is fairly low, approximately 34.14 years old, which unfortunately means that we may be limited in our abilities to detect heterogeneous effects related to age.
- *Health Checkup Frequency*: Individuals who prioritize regular health checkups might be more likely to respond better to the intervention. The response to the question, "*How often do you get a health checkup?*" may also reveal individuals with chronic health conditions that require them to make more frequent health checkups (every three or six months) and it is possible that their behaviors may differ from a generally healthy population. Therefore, it would be relevant to compare the subgroups of people who have a health checkup "Once a year" and "Only when necessary".
- *Vaccine Beliefs*: We asked the participants to answer the following question using a 11-point scale, "*Would you agree that it is important for everyone to get an influenza (flu) vaccine each year?*". Using a Net Promoter Score®, we found 141 *Detractors*, 87 *Passives*, and 163 *Promoters*, and we tested for heterogeneous treatment effects across a linear representation of this covariate.

In summary, we consider individuals who meet at least one (1) of the below criteria to have a high propensity for preventive healthcare:

- *Age* ≥ 40 years old
- *Health Checkup Frequency* = "Once a year"
- *Vaccine Belief* ≥ 7

## 3 Experimental Design

### 3.1 Study Setting and Data Collection Mechanism(s)

Our study was administered through a Qualtrics survey and the distribution of the intervention as well as the randomization of the participants across the treatment and control groups was accomplished through the Qualtrics platform. Data indicates that participants spent approximately *2.5 minutes*, on average, to complete the survey. Raw survey data was downloaded through the Qualtrics interface (see Appendix E to download a copy of the data).

### 3.2 Participants

We engaged participants through the Prolific Platform, which provides integration with Qualtrics. Adult individuals currently residing in the United States were invited to participate (we used a Prolific custom pre-screening option). Our survey was launched on Tuesday, March 17th, 2020, and we received a total of 391 completed survey responses.

Participants began the experiment by responding to some basic demographic questions, as well as a brief survey about their health practices and beliefs (see Appendix D for survey questions). The purpose of these background questions was to verify the diversity of the participant pool, control for key covariates, and test for heterogeneous treatment effects.

### 3.3 Intervention

Participants assigned to receive treatment were exposed to one of two different variations of the treatment in an attempt to assess if any one intervention produced different experimental outcomes (see Appendix D for CONSORT diagram). These are briefly described below.

#### 3.3.1 Treatment 1: Influenza-awareness video

Participants assigned to this treatment variation were presented with an informational, one-minute video that raises awareness about the dangers of influenza (flu) and preventive actions, including receiving a flu vaccine every year. The video is sponsored and hosted by the CDC, accessible [here](#). We enabled the collection of timing metrics that would identify non-compliers by virtue of how long participants spent on the page before navigating to the next screen. We found that, on average, participants spent 83.67 seconds on the page, and that approximately 7% of the participants (9 out of 130) did not view the entirety of the video. We decided not to categorize them as non-compliers, given that each of them received partial treatment. This highlights that in future experiments it might be advisable to introduce more rigorous controls in the survey flow, for example, highlighting the *Next* button only after delivery of the entire message.

#### 3.3.2 Treatment 2: Influenza-awareness quiz

Participants assigned to this treatment variation were presented with a seven-question multiple-choice quiz (see Appendix D) to test the individual’s knowledge about the flu, common misconceptions, and basic flu statistics. This intervention was inspired by the *Motivational Interviewing* approach that uses an “evoke and provide” strategy to elicit behavioral change. We hoped that this treatment might be more effective, especially for individuals with a low propensity for preventive healthcare. Participants were presented with their score at the end of the quiz, and the average score was approximately 4.8 points (out of 7). It is interesting to note that approximately 25% of the respondents believed that one can get the flu from the flu shot, which is a common misconception thought to be one of the main reasons behind low vaccination rates.

### 3.3.3 Control: Placebo message

Participants assigned to the Control group were simply presented a placebo message (see Appendix D) in anticipation of any possible bias introduced as a result of not experiencing a delay prior to outcome measurement (versus an anticipated 1-2 minute delay for participants assigned to treatment).

## 3.4 Outcome Measures

In an ideal world, we would have been able to measure an individuals' own flu vaccination propensity as the outcome variable. However, the presence of biases like *Social Desirability Bias* (where individuals have a tendency to over-report positive health behaviors and under-report undesirable ones) increases the likelihood of overestimating the true treatment effect. We instead decided to elicit an outcome response in the form of the participant's willingness to make a contribution to raising infectious disease awareness in the community.

Participants were presented with the choice of allocating a \$1 donation, with the option to divide the dollar as they desired between the National Foundation of Infectious Diseases (NFID) and the National Alliance on Mental Health (NAMI). Thus, the treatment effect was estimated by comparing the amount allocated to NFID among individuals in the treatment group(s) versus the control group. Figure 2 below shows the donation slider presented to each subject at the conclusion of the survey. Subjects were required to allocate the full donation before proceeding to the next screen.

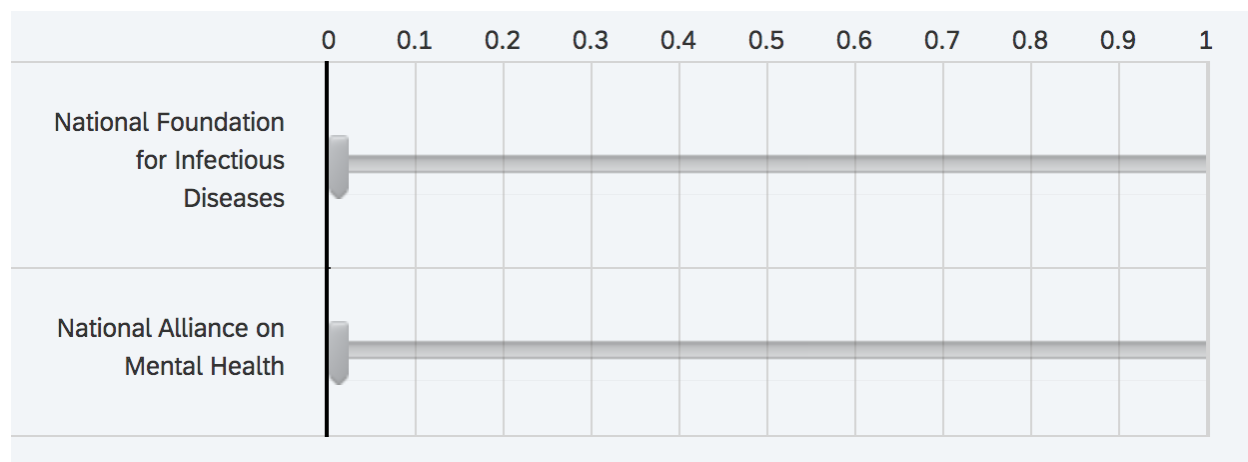


Fig. 2: Outcome measurement of dividing a 1 dollar donation between NFID and NAMI

We observed that approximately 39% of the respondents decided to split the donation equally between the two above causes. This highlights that there may be unobserved systemic reasons that may be related to an individuals' altruistic decisions, and in this case reflecting the uncertainty about which ways of doing good will actually be the most effective. The sum of all contributions was calculated to be \$197.57 towards NFID and \$193.43 towards NAMI, and the stipulated amounts have been donated by the research team to the respective charities.

## 3.5 Randomization

We used the Qualtrics infrastructure to perform simple random assignment to allocate the participants equally across the treatment and control group(s). This guaranteed that each participant had the same probability of being treated, and therefore treatment status should be statistically independent of the participants' potential outcomes and background attributes. This is evaluated through a table of means (see Table 1) and we were unable to reject the null hypothesis of no significant difference in means for each covariate across all experimental groups, using p-values obtained from an ANOVA test.

Table 1: Demographic and General Health Characteristics across Treatment and Control

	Table of Means			
	Control (131)	Video Treatment(130)	Quiz Treatment (130)	p-value (aov)
Age	32.99 (11.43)	34.92 (13.53)	34.52 (11.06)	0.40
Female	0.54	0.49	0.60	0.22
Nonwhite	0.34	0.28	0.33	0.53
High School degree or equivalent	0.35	0.37	0.32	0.65
General health index	7.16 (1.59)	7.19 (1.69)	7.47 (1.55)	0.24
Health insurance	0.82	0.88	0.87	0.43
Annual health checkup	0.42	0.30	0.39	0.11
Vaccine belief	7.25 (2.78)	6.86 (3.11)	7.34 (2.70)	0.36

Notes:  
 Num observations in column headers  
 Standard deviations in parenthesis

We also ran regression models of the treatment indicator variables on all pre-treatment covariates in order to confirm that the randomization effectively balanced covariates across the treatment and control groups (see Appendix B for model summary). The results of regressing the **Video Treatment** indicator on all covariates demonstrated that no covariate coefficients were statistically significant at the 0.05 level, indicating that the difference in means for all covariates between the video treatment and control were not statistically significant. Additionally, we could not reject the null hypothesis that the covariates do not jointly predict video treatment assignment (p-value=0.25).

We ran an analogous regression for the **Quiz Treatment** indicator across all covariates. The results of this model demonstrated that no covariate coefficients were statistically significant at the 0.05 level, and here again we could not reject the null hypothesis that the covariates do not jointly predict quiz treatment assignment (p-value=0.69).

These results suggest that the randomization process achieved covariate balance across the treatment and control groups.

### 3.6 Power Analysis

We conducted an a priori power analysis to identify the desired sample size and ensure the power was high enough that we would likely detect the assumed effect size. To perform the power analysis we simulated a schedule of potential outcomes under the following assumptions: (1) The distribution of contributions allocated to NFID in the control group is normally distributed with a mean of ~0.5 and a standard deviation of ~0.2, and (2) The effect of the quiz and video treatments on NFID contribution is approximately +\$0.05. With the simulated data, we then investigated the relationship between the experiment’s power and sample size (see Figure 4 below). Based on the power analysis, we determined that a sample size of approximately 400 participants would yield a power of ~0.70.

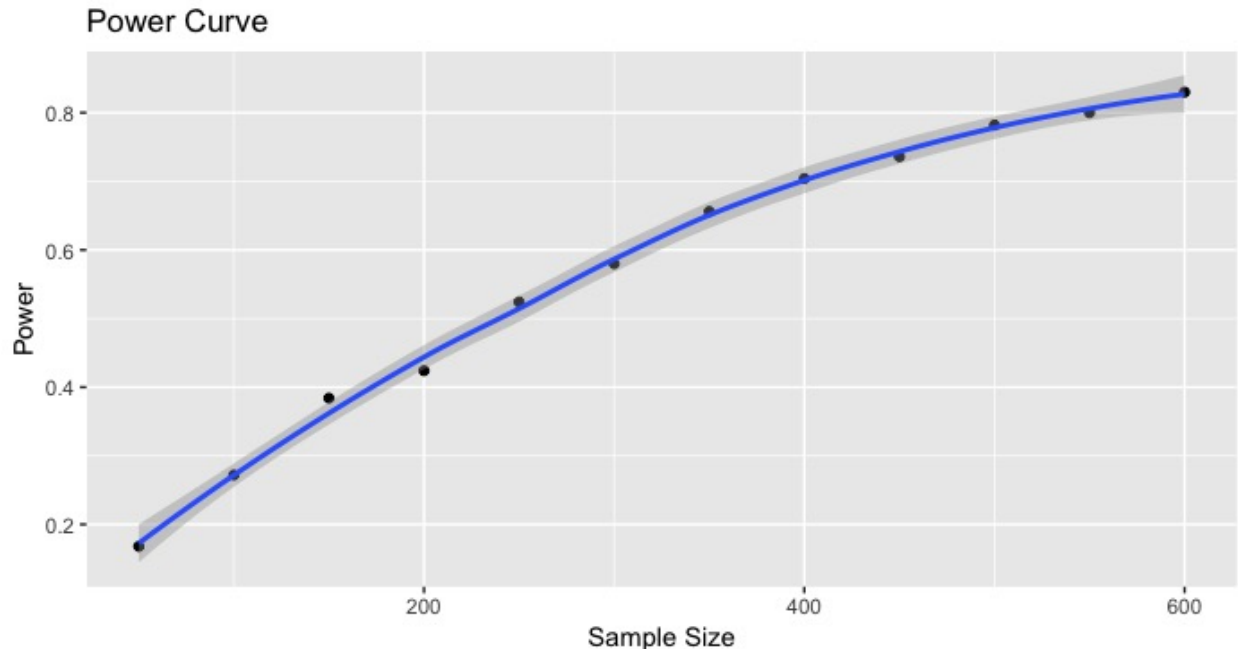


Fig. 3: Power as a function of experiment sample size

### 3.7 Pilot

We ran a pilot study from March 4th, 2020 through March 17th, 2020, where adult participants were acquired through the research teams’ community of family and friends. We received a total of 57 survey responses. The resulting table of means (see Appendix C) is evidence that all covariates except perhaps **Vaccine Belief** are well balanced across all experimental groups. Given the small sample size, the difference in means for **Vaccine Belief** may be attributed to high sampling variability.

We did find that roughly 14% of the respondents failed to provide an outcome measurement. As a result, we decided to incentivize the survey through a platform like Prolific, where respondents are paid a nominal fee for their participation. Additionally, we introduced stricter validation in the survey flow, to prevent missing data.

## 4 Results

### 4.1 Exploratory Data Analysis

The gender breakdown of participants was 54% female : 43% male, which is not perfect, but does represent a large enough sample of each to provide significant insights after accounting for the not-quite equal breakdown. Racially, participants were 69% White, 9% African American, 12% Asian American, and 7% Hispanic, which compares well to national averages<sup>7</sup>. Participants also covered wide range of educational backgrounds, with sizable portions all having High School, Bachelor’s, or Master’s Degrees. The vast majority had healthcare (86%), only slightly less than the national figure of 90%. Please see Appendix A for a summary of demographic and health distributions.

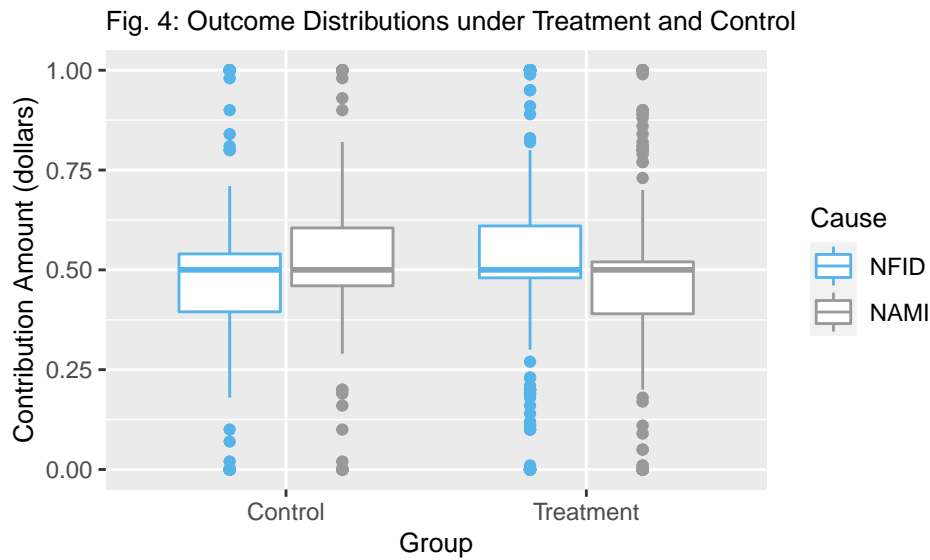
Out of the 391 participants, only 20 spent more than five minutes on the survey, and 39 spent less than one minute; both good indicators that most participants dedicated the appropriate amount of time and thus more likely they completed the survey honestly. No action was taken as a result of this observation. Only three participants seem to have mis-entered their birth years, two of whom entered their ages and another

who seemingly double-pressed one digit. For our analysis, we corrected these to the presumed birth years accordingly.

The most unfortunate issue in the data was simply a result of the limitations of the sample size. In exploring key relationships tied to the **Health Checkup Frequency** variable (a particular area of interest in our analysis), it became clear that only two response levels covered a large enough sample size to be reliably significant: “Once a year” and “Only when necessary”, each garnering 37% of the participants. Finally, we note that only 27% of the respondents were over 40 years old. Our initial hypothesis was that the treatment effect would be larger for older individuals compared to their younger counterparts. However, this distribution unfortunately means that we may be limited in detecting a heterogeneous effect related to age.

## 4.2 Distribution of Outcomes

Figure 4 below is a distribution of the outcome measures across the treatment (video and quiz combined) and control groups. Visual inspection of the box plots for each of the outcome measures, namely contributions to NFID and NAMI, reveal that they have very similar median values and variability across both treatment and control. However, one can observe that the interquartile range for NFID sits well below that of NAMI in the control group, and this reverses in the treatment group. *This is indicative of an apparent shift in the contribution trend (by participants) from control to treatment.*



## 4.3 Average Treatment Estimate

Table 2 reports the difference in means treatment effect estimates obtained as a result of regressing the outcome variable of interest on the **Video Treatment** and **Quiz Treatment** indicators, followed by a second model that includes a set of pre-treatment covariates that improve the precision of the treatment effect estimate(s). We identified these covariates using a stepwise regression approach which adds covariates that give the single best improvement to the model, as measured by the p-value from an ANOVA test at every step (forward). The forward steps were halted once the inclusion of each remaining covariate no longer gave a significant result from the ANOVA test at the 0.05 level. This process revealed that **Vaccine Belief**, **Gender**, **General Health Index**, and **Age** were the controls that best predicted the outcome. The standard errors presented are heteroskedastic robust errors.

Table 2: Treatment effect estimates of Raising Awareness on Promotion of Preventive Healthcare

	<i>Dependent variable:</i>	
	Contribution to National Foundation of Infectious Diseases	
	(1)	(2)
Video Treatment	0.009 (0.035)	0.014 (0.033)
Quiz Treatment	0.037 (0.032)	0.026 (0.031)
Vaccine belief		0.022*** (0.005)
Male (gender)		0.060** (0.027)
Non-Binary (gender)		-0.189* (0.105)
General health index		0.024*** (0.009)
Age		0.003** (0.001)
Constant	0.490*** (0.023)	0.040 (0.079)
Observations	391	391
R <sup>2</sup>	0.003	0.121
Adjusted R <sup>2</sup>	-0.002	0.105
Residual Std. Error	0.270 (df = 388)	0.255 (df = 383)
F Statistic	0.652 (df = 2; 388)	7.505*** (df = 7; 383)
<i>Note:</i>		*p<0.1; **p<0.05; ***p<0.01 Robust Standard Errors

The regression of the outcome variable of interest, NFID contribution amount, on the **Video Treatment** and **Quiz Treatment** indicator variables shows that the estimated average treatment effect for the **Video Treatment** is an increase of **0.9¢**, while the average treatment effect for the **Quiz Treatment** is an increase of **3.7¢**. However, both of these estimates are statistically insignificant at the 0.05 level, and thus these effects may be attributable to chance.

Controlling for covariates, the estimated average treatment effect of the **Video Treatment** is an increase of **1.4¢** and the average treatment effect of the **Quiz Treatment** is an increase of **2.6¢**. All pre-treatment covariates have statistically significant coefficients at the 0.10 level. We also note that while the estimates did not change drastically with the inclusion of these covariates, the standard errors of both the **Video Treatment** and **Quiz Treatment** estimates decreased, re-affirming that the covariates included in the model are “good controls”. However, it is still important to note that both of the treatment estimates remained statistically insignificant at the 0.05 level.

#### 4.4 Are certain individuals more responsive to the intervention?

In an effort to test our hypothesis that individuals with a higher propensity for preventive healthcare will be more responsive to the intervention, Table 3 reports the heterogeneous treatment effects obtained as a result of regressing the outcome variable on the **Video Treatment** and **Quiz Treatment** variables, and interaction terms between the treatment and covariates of interest (**Age**, **Health Checkup Frequency**, and **Vaccine Belief**), in three separate models. The standard errors presented are heteroskedastic robust errors.

Next, we investigate the statistical significance of the treatment variables along with the interaction terms.



Table 3: Heterogenous Treatment effect estimates by Age, Health Checkup Frequency, and Vaccine Belief

	<i>Dependent variable:</i>		
	Contribution to National Foundation of Infectious Diseases		
	Age (Binned)	Health Checkup Frequency	Vaccine Belief
	(1)	(2)	(3)
Video Treatment	0.005 (0.037)	−0.067 (0.060)	−0.189** (0.095)
Quiz Treatment	0.047 (0.035)	−0.024 (0.056)	0.062 (0.106)
Vaccine belief	0.022*** (0.005)	0.023*** (0.005)	0.012 (0.009)
Male (gender)	0.057** (0.027)	0.042 (0.027)	0.062** (0.026)
Non-Binary (gender)	−0.199* (0.106)	−0.215* (0.112)	−0.192* (0.102)
General health index	0.024*** (0.009)	0.023** (0.009)	0.024*** (0.009)
Health checkup once a year		−0.096* (0.056)	
Health checkup once in six months		−0.028 (0.072)	
Health checkup once in three months		−0.224 (0.149)	
Health checkup never		−0.101 (0.075)	
Age		0.003** (0.001)	0.003** (0.001)
Video:Health checkup once a year		0.164** (0.081)	
Video:Health checkup once in six months		0.004 (0.096)	
Video:Health checkup once in three months		0.191 (0.186)	
Video:Health checkup never		0.231 (0.194)	
Quiz:Health checkup once a year		0.046 (0.074)	
Quiz:Health checkup once in six months		0.017 (0.091)	
Quiz:Health checkup once in three months		0.168 (0.177)	
Quiz:Health checkup never		0.353*** (0.111)	
Video:Vaccine belief			0.029** (0.012)
Quiz:Vaccine belief			−0.005 (0.013)
Age >= 40	0.075 (0.057)		
Video:Age >= 40	0.023 (0.081)		
Quiz:Age >= 40	−0.078 (0.076)		
Constant	0.123 (0.076)	0.119 (0.091)	0.115 (0.099)
Observations	391	391	391
R <sup>2</sup>	0.119	0.160	0.147
Adjusted R <sup>2</sup>	0.098	0.117	0.127
Residual Std. Error	0.256 (df = 381)	0.254 (df = 371)	0.252 (df = 381)
F Statistic	5.692*** (df = 9; 381)	3.711*** (df = 19; 371)	7.300*** (df = 9; 381)

Note:

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

Robust Standard Errors

Reference level for Age: Age &lt;40 years

Reference level for Health Checkup Frequency: Only when necessary

#### 4.4.1 Treatment Effect Heterogeneity by Age

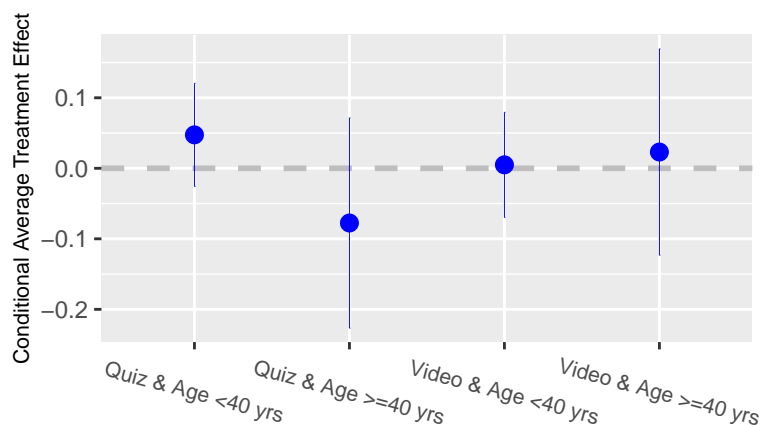


Fig. 5: Coefficient Plot of Heterogeneous Treatment Effects by Age. Each point represents estimates from an OLS regression, with 95 percent confidence intervals.

Results of the regression that evaluates the heterogeneous treatment effects of **Age** show that the estimated treatment effect of the **Video Treatment** is **2.3¢** greater for individuals  $\geq 40$  years old versus those  $< 40$  years old. We hypothesized that the treatment effect would be greater for individuals  $\geq 40$  years old, and thus this estimate is consistent with our hypothesis. However, the heterogeneous treatment effect is statistically insignificant.

Conversely, the estimated treatment effect of the **Quiz Treatment** is **7.8¢** less for individuals  $\geq 40$  years old versus those  $< 40$  years old. This estimate is also statistically insignificant.

#### 4.4.2 Treatment Effect Heterogeneity by Health Checkup Frequency

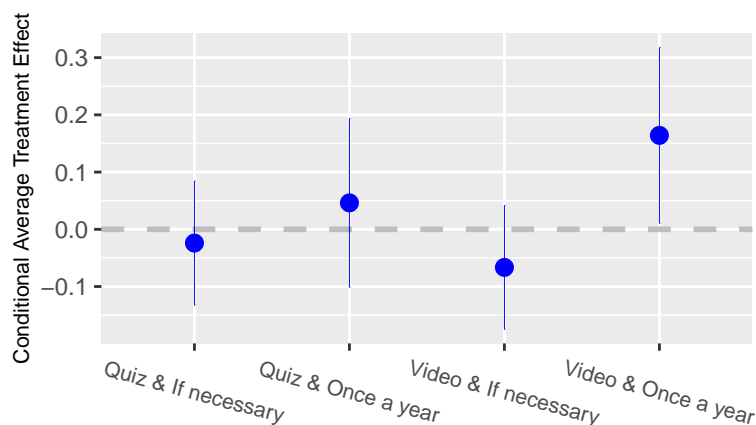


Fig. 6: Coefficient Plot of Heterogeneous Treatment Effects by Health Checkup Frequency. Each point represents estimates from an OLS regression, with 95 percent confidence intervals.

Results of the regression that evaluates the heterogeneous treatment effects of **Health Checkup Frequency** show some interesting relationships. First, the estimated treatment effect of the **Video Treatment** is **-6.7¢**,

but is not statistically significant at the 0.05 level. Since we have included interaction terms in this model, this coefficient represents the estimated treatment effect for individuals in the base **Health Checkup Frequency** level of “Only when necessary”. Further, the coefficient of the interaction term between the **Video Treatment** and “Once a year” **Health Checkup Frequency** is statistically significant at the 0.05 level. This suggests that the **Video Treatment** is **16¢** greater for individuals with a higher propensity for preventive health (“Once a year” **Health Checkup Frequency**) relative to individuals with a lower propensity (“Only when necessary” **Health Checkup Frequency**). Finally, it is also interesting to note that while the conditional average treatment effect of the video is positive for individuals in the “Once a year” group, the effect is negative for individuals in the “Only when necessary” and “Once in 6 months” groups.

The treatment effect heterogeneity for the **Quiz Treatment** is statistically significant for individuals in the “Never” **Health Checkup Frequency** group at the 0.01 level. This suggests that the quiz treatment effect for individuals who never have physician checkups is **35¢** greater than the quiz treatment effect for individuals who have checkups only when necessary. However, only 24 individuals fall within the “Never” **Health Checkup Frequency** group, so we remain skeptical that this estimate may be the result of chance.

#### 4.4.3 Treatment Effect Heterogeneity by Vaccine Belief

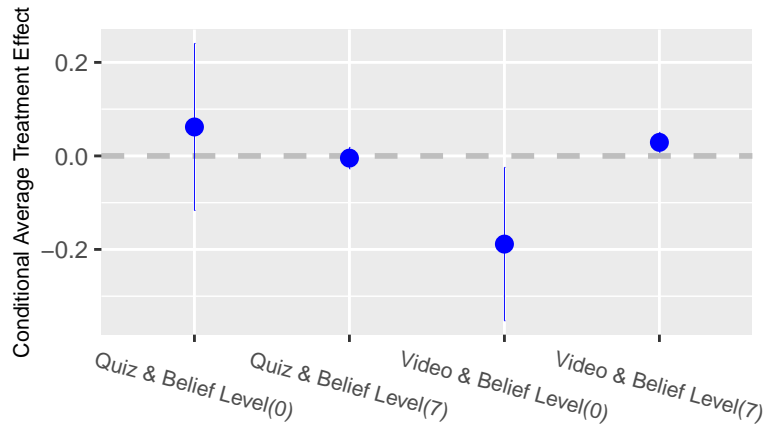


Fig. 7: Coefficient Plot of Heterogeneous Treatment Effects by Vaccine Belief. Each point represents estimates from an OLS regression, with 95 percent confidence intervals.

Results of this regression demonstrate a statistically significant heterogeneous treatment effect between the **Video Treatment** and **Vaccine Belief** variable. The coefficient of the **Video Treatment** indicator variable is **-0.19¢**, which represents the **Video Treatment** effect for individuals with a **Vaccine Belief** of 0 (i.e. individuals who strongly oppose the vaccine). This coefficient is statistically significant at the 0.05 level. The coefficient of the **Video Treatment** and **Vaccine Belief** interaction term suggests that for every 1-point increase in the **Vaccine Belief** scale, the treatment effect of the video increases by approximately **3¢**. This heterogeneous treatment effect is also statistically significant at the 0.05 level. The combination of these two estimates demonstrates that the **Video Treatment** is negative for individuals with a **Vaccine Belief** under 7 and positive for individuals with a **Vaccine Belief** of 7 and above. Lastly, neither the treatment effect of the quiz nor its heterogeneous effect with **Vaccine Belief** are statistically significant.

We performed an ANOVA test for each of 3 above models against the difference in means regression model controlling for covariates. The **Age** and **Health Checkup Frequency** models produced statistically insignificant results. However, the **Vaccine Belief** model produced a statistically significant result at the 0.01 level (p-value=0.003). This suggests that the inclusion of the **Vaccine Belief** heterogeneous treatment effects did improve overall performance from the base model.

#### 4.4.4 Adjusting for Multiple Comparisons

A practical implication of testing for multiple significant interactions is to adjust the p-value(s) accordingly to guard against false discoveries. We used the false discovery rate (`fdr`) method of correction (Benjamini & Hochberg<sup>6</sup>) which is a less conservative correction than the family-wise error rate (`bonferroni`, for example). Since we are testing our hypotheses across 12 different interactions of the treatment and covariate levels (`Age(2)`, `Health Checkup Frequency(8)`, and `Vaccine Belief(2)`), we ran the p-values through the appropriate `p.adjust()` method and found that the previously highly significant interactions of `Quiz:Health checkup` never saw a p-value adjustment from **0.0016 to 0.0195**, for `Video:Vaccine belief` from **0.0159 to 0.0951**, and for `Video:Health checkup once a year` from **0.0429 to 0.1715**. The first 2 interactions continue to be significant at the 0.1 level.

## 5 Discussion and Conclusion

Although the two estimated treatment effects did not appear to be statistically significant, the analysis of the treatment effect heterogeneity did reveal some statistically significant effects that support our initial hypothesis. First, we note that the `Video Treatment` effect for individuals with a `Health Checkup Frequency` of “Once a year”, representing individuals with a higher propensity to take preventive health measures, was **16¢** greater than individuals with a `Health Checkup Frequency` of “Only when necessary”, representing individuals with a lower propensity. Furthermore, the coefficient of the interaction term between the `Vaccine Belief` covariate and `Video Treatment` variable indicated that for each 1-unit increase in `Vaccine Belief`, the `Video Treatment` effect increased by **3¢**. The `Video Treatment` effect was positive for individuals with a `Vaccine Belief` greater than or equal to 7 and negative for individuals with a `Vaccine Belief` below 7. The combination of these relationships directly supports our initial hypothesis that individuals with a higher propensity to take preventive health measures would have a stronger response to the treatment.

While we identified a statistically significant effect between the “Never” `Health Checkup Frequency` group and `Quiz Treatment`, we are skeptical that we can draw conclusions from this estimate. We unfortunately did not identify any other statistically significant effect with respect to the `Quiz Treatment`. We also did not identify a statistically significant heterogeneous treatment effect for the `Age` covariate. However, our ability to detect this heterogeneous treatment effect may have been deterred by the skewed distribution with respect to `Age`.

It is also important to note that this study took place during the COVID-19 pandemic. While it is difficult to measure the direction of the bias stemming from the timing of the study, one possibility is that individuals with a higher propensity for preventive health measures may have had a positively biased response to treatment, given the abundance of news stories and cautions circulating through the media.

In conclusion, the results of the above analysis ultimately demonstrate that raising awareness did cause health-conscious individuals to promote the prevention of infectious diseases.

## Acknowledgements

We thank D. Alex Hughes for his guidance in designing and conducting the experiment. The experiment was funded by the UC Berkeley School of Information.

## References

1. “Up to 650 000 People Die of Respiratory Diseases Linked to Seasonal Flu Each Year.” World Health Organization, World Health Organization, 14 Dec. 2017, <http://www.who.int/news-room/detail/14-12-2017-up-to-650-000-people-die-of-respiratory-diseases-linked-to-seasonal-flu-each-year>.

2. Somes, Mitchell P., et al. "Estimating the Annual Attack Rate of Seasonal Influenza among Unvaccinated Individuals: A Systematic Review and Meta-Analysis." *Vaccine*, Elsevier, 30 Apr. 2018, <http://www.sciencedirect.com/science/article/pii/S0264410X18305607?via%3Dihub>.
3. "Flu Vaccination Coverage, United States, 2018–19 Influenza Season." Centers for Disease Control and Prevention, Centers for Disease Control and Prevention, 26 Sept. 2019, <http://www.cdc.gov/flu/fluview/coverage-1819estimates.htm>.
4. Young, Eric. "41 Percent of Americans Do Not Intend to Get a Flu Shot This Season." NORC at the University of Chicago, 5 Dec. 2018, <http://www.norc.org/NewsEventsPublications/PressReleases/Pages/41-percent-of-americans-do-not-intend-to-get-a-flu-shot.aspx>.
5. Mitchell, David. "Survey Reveals Common Misconceptions About Flu, Vaccination." AAFP Home, American Academy of Family Physicians (AAFP), 22 Jan. 2020, <http://www.aafp.org/news/health-of-the-public/20200122flusurvey.html>.
6. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing
7. "U.S. Census Bureau QuickFacts: United States." Census Bureau QuickFacts, <http://www.census.gov/quickfacts/fact/table/US/PST045218>.

# Appendix

## Appendix A: Exploratory Data Analysis

Fig 8: Demographic Distributions of Participants

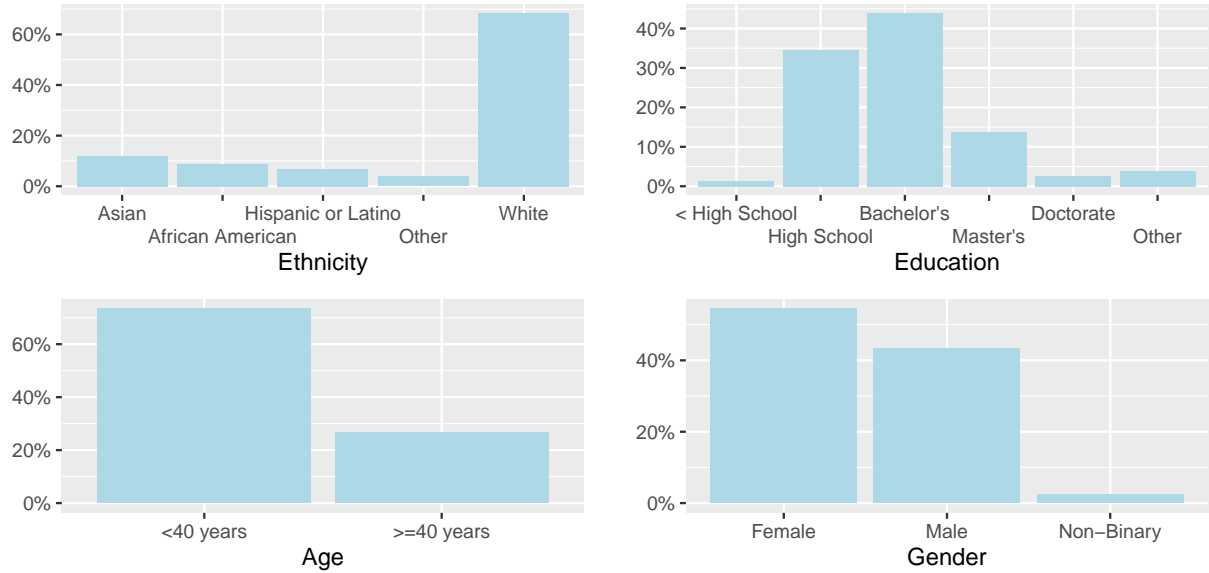
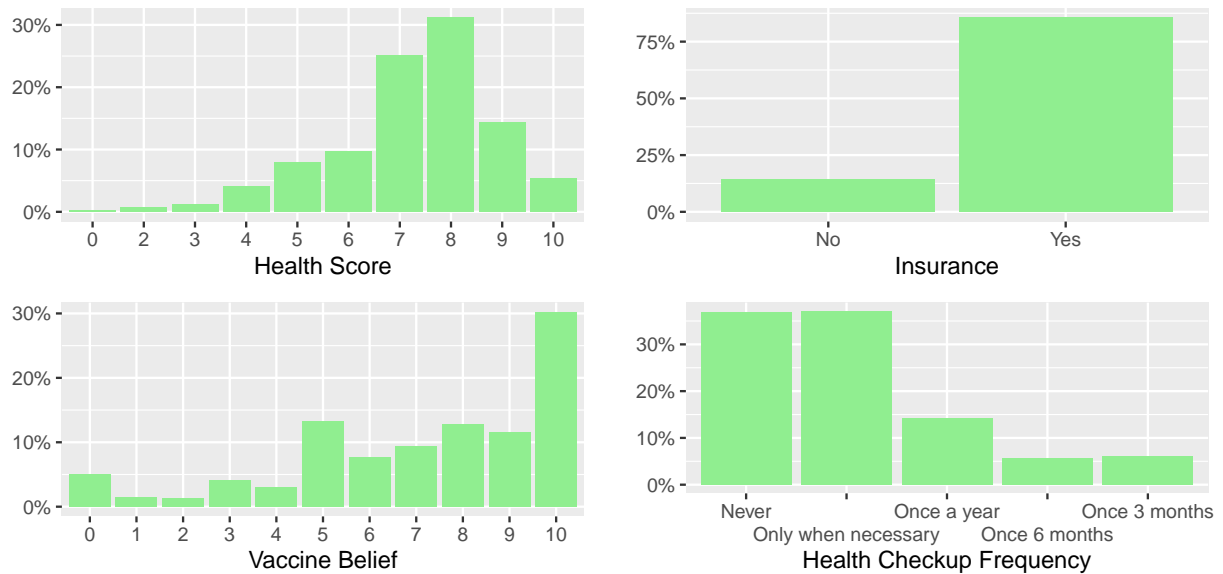


Fig 9: General Health Distributions of Participants



## Appendix B: Models

### Covariate Balance using Regression

Table 4: Covariate Balance Checks using Regression

	<i>Dependent variable:</i>	
	Video Treatment	Quiz Treatment
	(1)	(2)
Age	0.002 (0.002)	0.001 (0.002)
Black or African American	-0.152 (0.113)	0.112 (0.112)
Hispanic or Latino	-0.071 (0.113)	0.059 (0.116)
Other (ethnicity)	0.045 (0.158)	0.171 (0.157)
White	-0.016 (0.080)	0.019 (0.075)
Male	0.067 (0.051)	-0.065 (0.050)
Non-Binary	0.273 (0.209)	-0.157 (0.172)
High School degree or equivalent	-0.251 (0.265)	0.066 (0.212)
Bachelor's degree	-0.315 (0.265)	0.067 (0.213)
Master's degree	-0.183 (0.273)	0.106 (0.220)
Doctorate	-0.205 (0.309)	0.223 (0.275)
Other (education)	-0.301 (0.298)	0.137 (0.247)
General health index	-0.007 (0.017)	0.026* (0.016)
Health checkup only when necessary	0.153 (0.103)	0.056 (0.109)
Health checkup once a year	0.056 (0.104)	0.026 (0.109)
Health checkup once in six months	0.186 (0.119)	-0.085 (0.115)
Health checkup once in three months	0.186 (0.149)	0.085 (0.148)
Vaccine belief	-0.011 (0.009)	0.010 (0.009)
Constant	0.534* (0.311)	-0.052 (0.251)
Observations	391	391
R <sup>2</sup>	0.055	0.038
Adjusted R <sup>2</sup>	0.009	-0.009
Residual Std. Error (df = 372)	0.469	0.474
F Statistic (df = 18; 372)	1.208	0.813

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01  
Robust Standard Errors

## Appendix C: Pilot Experiment

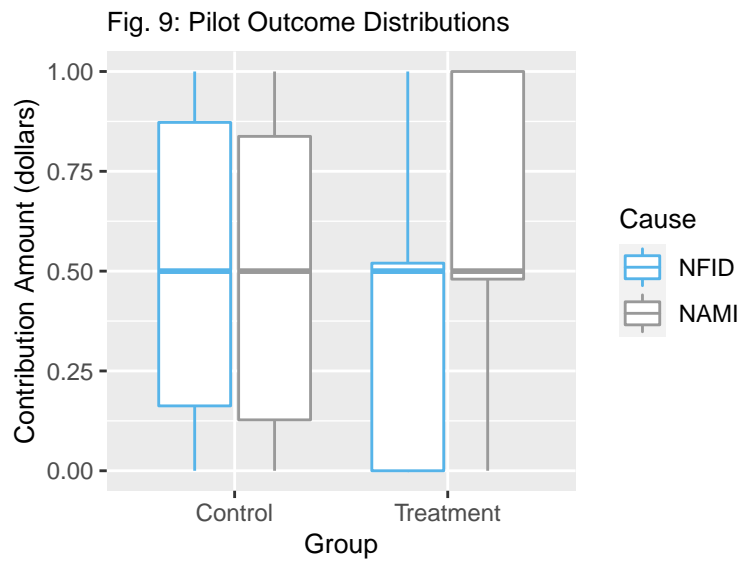
Table 5: Pilot Covariate Balance Check across Treatment and Control

	Table of Means			p-value (aov)
	Control (19)	Video Treatment(20)	Quiz Treatment (18)	
Age	33.14 (16.95)	30.06 (11.76)	28.81 (9.48)	0.64
Female	0.29	0.56	0.50	0.30
Nonwhite	0.64	0.39	0.56	0.35
High School degree or equivalent	0.07	0.17	0.19	0.65
General health index	8.00 (1.11)	7.44 (1.29)	7.88 (1.02)	0.36
Health insurance	0.86	0.94	1.00	0.28
Annual health checkup	0.71	0.56	0.44	0.33
Vaccine belief	5.86 (3.82)	8.61 (2.06)	8.44 (1.93)	0.01

Notes:

Num observations in column headers

Standard deviations in brackets





## Appendix D: Survey CONSORT Diagram and Questions

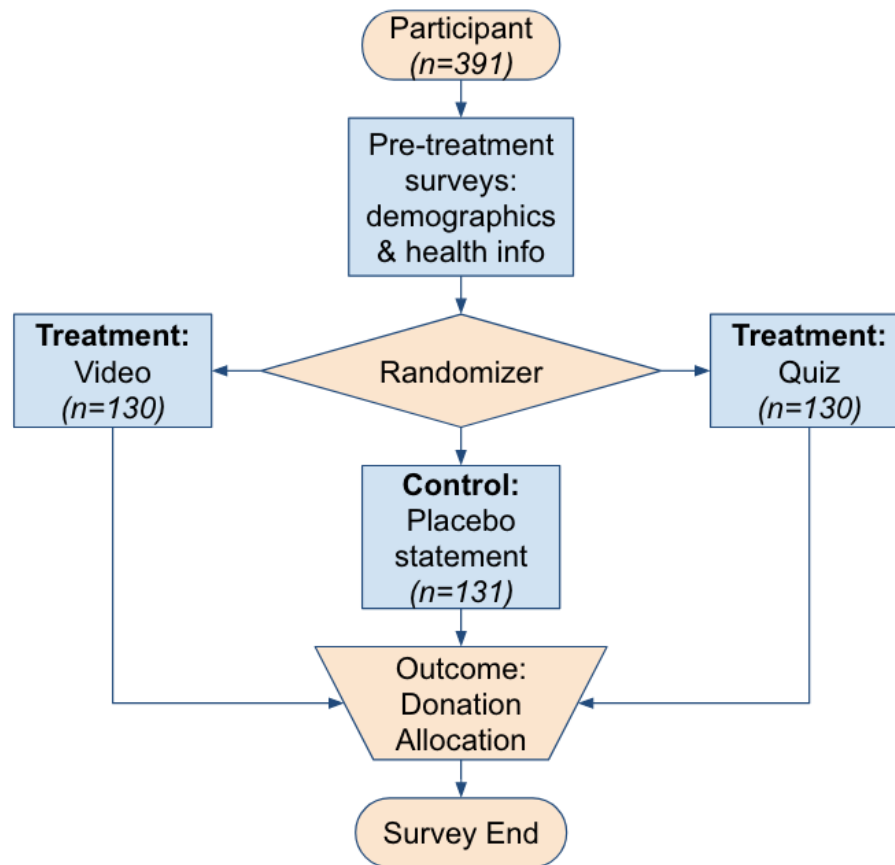


Fig. 10: CONSORT Diagram tracking the flow of participants during the experiment

### Demographic Survey

1. Please specify your year of birth (free-form text)
2. What gender do you identify as?
  - Male
  - Female
  - Non-Binary
  - Prefer to self-describe
  - Prefer not to answer
3. Please specify your ethnicity
  - American Indian or Alaska Native
  - Asian
  - Black or African American
  - Hispanic or Latino
  - Native Hawaiian or Pacific Islander
  - White
  - Other (please specify)

- Prefer not to answer
4. What is the highest degree or level of education you have completed?
- Less than a high school diploma
  - High School degree or equivalent
  - Bachelor's degree
  - Master's degree
  - Doctorate
  - Other (please specify)
  - Prefer not to answer

### **Health Survey**

1. How would you rate your health in general?
- **Poor to Excellent** on a 0-10 scale
2. Do you currently have health insurance?
- Yes
  - No
3. How often do you get a health checkup?
- Once in 3 months
  - Once in 6 months
  - Once a year
  - Only when necessary
  - Never
4. Would you agree that it is important for everyone to get an influenza (flu) vaccine each year?
- **Strongly Disagree to Strongly Agree** on a 0-10 scale

### **Quiz Treatment**

1. Influenza is caused by
- A type of bacteria
  - A virus
  - A parasitic organism
  - I don't know what causes influenza
2. How does the flu spread?
- Through spoiled food
  - Through the air or by touching an infected surface
  - Through blood
3. What chance does a healthy person have of getting the flu during an average year?
- Between 70 and 80 percent
  - Between 40 and 50 percent
  - Between 10 and 20 percent
  - Between 0 and 2 percent
4. How long are you contagious once you have the flu?
- 1 day before symptoms start

- 1 day before symptoms start and up to 1 week after becoming sick
  - 3-4 days after becoming sick
  - 1 week after becoming sick
5. You can get the flu from a flu shot.
- True
  - False
6. What is the best time of year for a person to be vaccinated against influenza?
- In autumn, just before the peak flu season
  - In summer, well before the next peak flu season
  - As soon as flu symptoms develop
  - Any time; it makes no difference
7. How many hospitalizations were caused by the flu in 2018-2019?
- 30,000
  - 100,000
  - 500,000
  - 1,000,000

### **Placebo Message**

Donating to the causes you care about not only benefits the charities themselves, but it can be deeply rewarding for you too. Millions of people give to charity on a regular basis to support causes they believe in, as well as for the positive effect it has on their own lives.

### **Outcome Measurement Message**

How would you like us to allocate a \$1 donation to the following non-profits on your behalf?

- The National Foundation for Infectious Diseases, an organization dedicated to educating the public and healthcare professionals about the burden, causes, prevention, diagnosis, and treatment of infectious diseases.
- NAMI, or National Alliance on Mental Health, an organization dedicated to improving the quality of life for people with mental illness and their families through support, education, and advocacy.

## **Appendix E: Code and Data**

Code and Data can be found at the following locations in our GitHub repository:

- Raw pilot data
- Raw experiment data
- Cleaned data
- Analysis code