Pre-Registration and Pre-Analysis Plan:
Public Perception of Ethical Trade-offs in COVID-19 Vaccine Trial Design

We will be conducting an online survey to gauge how the public weighs various ethical trade-offs related to the design of vaccine trials. We will be conducting this survey in the United States, United Kingdom, Australia, Canada, New Zealand, Singapore, Hong Kong, and South Africa using the online survey provider Lucid. We aim for a sample size of 500 respondents per country and 2,000 in the United States. Respondents must take the survey on a desktop and be able to read English.

In the survey, we first ask two attention check questions. If respondents fail those attention checks, they are removed from the survey.

We then randomly assign respondents to one of two studies:
- Challenge Trial vs. Standard Design
- Integrated Phase 2/3 vs. Standard Design

In the Challenge Trial vs. Standard Design study, we randomize the following features:
- Standard Design Trial N (3,000; 5,000; 7,000; 9,000; 11,000)
- Challenge Trial N (80; 100; 200)
- % in Standard Design that are exposed to coronavirus in their daily lives (2%, 5%, 20%)
- % in Standard Design who die of coronavirus if they are exposed (0.5%, 1%)
- How long the Standard Design takes to get a vaccine ready: 12 or 18 months
- How much faster the Challenge Trial is: 2, 4, or 6 months faster
- Which design is described as "Study A" or "Study B"

In the Integrated Phase 2/3 vs. Standard Design study, we randomize the following features:
- Standard Design Trial N (3,000; 5,000; 7,000; 9,000; 11,000)
- How long the Standard Design takes to get a vaccine ready: 12 or 18 months
- How much faster the Integrated Design is: 2, 4, or 6 months faster
- How many people in the Integrated Design are exposed to the vaccine before it stops, in the case that the vaccine in found to be unsafe (250; 400; 1,000)
- Which design is described as "Study A" or "Study B"

For each study, we will ask the following outcome measures:
- Primary outcome
  - If you had to choose, which study would you rather have scientists conduct? Study A; Study B
    - This will be coded as 1 for Challenge/Integrated and 0 for Standard
- Secondary outcomes
  - How ethical do you think the studies are? Asked for both designs
    - Definitely ethical (4); Probably ethical (3); Probably unethical (2); Definitely unethical (1)

- ■ We will analyze this outcome by taking the difference between Challenge/Integrated minus Standard
- ■ We will also report the frequencies for the individual variables
  - ○ How scientifically valid do you think the studies are? Asked for both Study A and Study B
    - ■ Very valid (4); Somewhat valid (3); Somewhat invalid (2); Very invalid (1)
    - ■ We will analyze this outcome by taking the difference between Challenge/Integrated minus Standard
    - ■ We will also report the frequencies for the individual variables
  - ○ If the study found the vaccine worked and it was then approved by the government, how likely would you be to take the vaccine to protect yourself from coronavirus? Asked for both Study A and Study B
    - ■ Very likely (4); Somewhat likely (3); Somewhat unlikely (2); Very unlikely (1)
    - ■ We will analyze this outcome by taking the difference between Challenge/Integrated minus Standard
    - ■ We will also report the frequencies for the individual variables

For each study, we will also ask the following factual understanding questions to ensure respondents, on average, paid attention and understood the survey:
- ● If the vaccine works, which of the two studies we asked about would result in the vaccine being approved and widely available sooner?
- ● If the vaccine works, which of the two studies we asked about would result in more people in society generally dying of coronavirus?
- ● These questions will only be asked for the Challenge Trial vs. Standard Design study:
  - ○ Which of the two studies we asked about involves intentionally exposing participants to coronavirus while they are quarantined in a medical research center?
  - ○ Which of the two studies we asked about would result in more people in the study dying of coronavirus?
    - ■ Note: This question has a "Neither" option because in some randomizations, the number is the same.
- ● This question will only be asked for the Integrated vs. Standard Design study:
  - ○ Which of the two studies we asked about involves doing additional safety testing on a smaller group first?
- ● Each factual understanding variable will be recoded to have 1 for the correct answer and 0 for the incorrect answer based on the randomization.

We will conduct the below analyses for each study (challenge and integrated). For each average we describe below, we will perform a one-sample t-test, testing the null hypothesis that the challenge/integrated and standard designs are equal; this implies a null of 0.5 for the "If you had to choose" variable and a null of 0 for the secondary outcomes.
- ● Average value for each outcome, overall and by country
- ● Subgroups of primary interest are listed below. Our goal for subgroup analyses is to demonstrate the consistency of the findings across a) randomized descriptions of trial designs and b) salient

social cleavages, especially among vulnerable populations and politically relevant groups. With this in mind, we will compute the average value and perform the t-tests mentioned above among participants in each of the subgroups mentioned below. We will only examine subgroups that are at least N=50 in size.

- Demographic groups
  - Only people 65 and over, given they are at highest risk for serious complications or death from coronavirus
  - Only participants who understood all the factual understanding questions correctly
  - Only participants who say they are "essential workers"
  - Racial minorities. We will measure this as follows:
    - US / UK / Australia / NZ / South Africa / Canada: those who do not select "White" to the race/ethnicity question
    - Singapore and HK: Those who do not select "Chinese" to a race/ethnicity question
  - Generally speaking, do you consider yourself a...Democrat/Republican/Independent/Other Party. We will create indicators for each, pooling Independents and Other Party into one category. (This analysis will be done for US respondents only.)
  - In a US county with cumulative COVID cases per capita above the median. To calculate COVID cases per capita, we will compute county population using 2019 Census population estimates (https://www2.census.gov/programs-surveys/popest/tables/2010-2019/counties/totals/co-est2019-annres.xlsx) and COVID cases determined on the date of the launch of the survey, using the *New York Times* data at https://raw.githubusercontent.com/nytimes/covid-19-data/master/us-counties.csv. (This analysis will be done for US respondents only.)
- Randomized descriptions of trial designs
  - By the number of months the Standard design takes and how much faster the Challenge/Integrated designs will be, as a 2x3 table with 6 separate statistics (reported separately for Integrated and Challenge)
  - Average value for each outcome by the death rate and sick rate in the trial, as a 2x3 table with 6 separate statistics (Challenge only)
  - Average value for each outcome by number of people who get the vaccine before it is determined to be unsafe (Integrated only)

- For each outcome, we will also report a regression to estimate which demographics predict support. We may also report raw means of outcomes within demographic categories. We will include the following predictors, all as linear predictors unless specified otherwise:
  - In your opinion, how important is it that parents get their children vaccinated? Extremely important (5); Very important (4); Somewhat important (3); Not very important (2); Not at all important (1)
  - Do you think vaccines are more dangerous than the diseases they are designed to prevent, or not? Yes (3); Unsure (2); No (1)

- How concerned are you about the effect of the coronavirus on the country's economy? Very concerned (4); Somewhat concerned (3); Not very concerned (2); Not at all concerned (1)
- How concerned are you that you, someone in your family, or someone else you know will become infected with coronavirus? Very concerned (4); Somewhat concerned (3); Not very concerned (2); Not at all concerned (1)
- What is your year of birth? Recoded as age and groups into bins: 18-24; 25-44; 45-64; 65+, each analyzed as an indicator variable
- Which of the following best describes your gender? 1 = female; 0 = all other
- What is the highest level of education you have completed? 1 = college educated or above; 0 = all other
- In political matters, people talk of "the left" or "liberal" and "the right" or "conservative". How would you place your views on this scale, generally speaking? Coded from 1 (Liberal) to 10 (Conservative)
- What is your current employment status, and are you considered an "essential worker" during this pandemic? Indicators for employed as an essential worker; employed as a non-essential worker; unemployed due to COVID; furloughed due to COVID; and retired.
- Apart from weddings and funerals, about how often did you usually attend religious services last year? More than once a week (5); Once a week (4); Once a month (3); Only on special holy days (2); Once (1); Never (0)
- Scientific knowledge. We ask respondents if they know "Which kind of waves is used to make and receive cellphone calls?" and "Ocean tides are created by which of the following?". Respondents get a 2 if they answer both correctly; a 1 if they answer one correctly; and a 0 if they answer none correctly
- An indicator for every country.
- Race/ethnicity: We will create indicators for US Black, US Asian, US Latino, and for non-white in each of the UK, Australia, NZ, South Africa, and Canada, and for non-Chinese in Singapore and Hong Kong.
- An indicator for whether the integrated/challenge study was randomized to be "Study A" or "Study B".

For a separate research question, we will also analyze the mean for the question examining preferences about a post-challenge trial safety study involving either 3,000 or 1 million people, both overall and among those 65+ only. We may report both these results separately.

We may also conduct qualitative analyses of the open-ended responses to the questions asking respondents why they gave the answers about the ethics they did. Details of how we will conduct this analysis are not pre-registered.

Our primary analyses will be unweighted. As a robustness check, we will also present results for the United States using weights. For this analysis, we will weight to the ACS on age, gender, education, and

race using the ebalance package in Stata. We will compute the weighted mean using the wtd.t.test function from the weights package in R. Our analysis will assume the weights are fixed.