

Broad Cross-National Public Support for Accelerated COVID-19 Vaccine Trial Designs

David Broockman* Joshua Kalla† Alexander Guerrero‡
Mark Budolfson§ Nir Eyal¶ Nicholas P. Jewell|| Monica Magalhaes**
Jasjeet S. Sekhon††

December 20, 2020

Abstract

A vaccine for COVID-19 is urgently needed. Several vaccine trial designs may significantly accelerate vaccine testing and approval, but also increase risks to human subjects. Concerns about whether the public would see such designs as ethical represent an important roadblock to their implementation; accordingly, both the World Health Organization and numerous scholars have called for consulting the public regarding them. We answered these calls by conducting a cross-national survey ($n = 5,920$) in Australia, Canada, Hong Kong, New Zealand, South Africa, Singapore, the United Kingdom, and the United States. The survey explained key differences between traditional vaccine trials and two accelerated designs: a challenge trial or a trial integrating a Phase II safety and immunogenicity trial into a larger Phase III efficacy trial. Respondents' answers to comprehension questions indicate that they largely understood the key differences and ethical trade-offs between the designs from our descriptions. We asked respondents whether they would prefer scientists to conduct traditional trials or one of these two accelerated designs. We found broad majorities prefer for scientists to conduct challenge trials (75%) and integrated trials (63%) over standard trials. Even as respondents acknowledged the risks, they perceived both accelerated trials as similarly ethical to standard trial designs. This high support is consistent across every geography and demographic subgroup we examined, including vulnerable populations. These findings may help assuage some of the concerns surrounding accelerated designs.

Keywords: COVID-19; vaccine ethics; challenge trials; public opinion

*Department of Political Science, University of California, Berkeley.

†To whom correspondence should be addressed: josh.kalla@yale.edu. Department of Political Science and Department of Statistics & Data Science, Yale University.

‡Department of Philosophy, Rutgers University.

§Department of Environmental and Occupational Health and Justice, Rutgers University.

¶Center for Population-Level Bioethics, Rutgers University.

||London School of Hygiene & Tropical Medicine and Division of Biostatistics, School of Public Health, University of California, Berkeley.

**Center for Population-Level Bioethics, Rutgers University.

††Department of Statistics & Data Science and Department of Political Science, Yale University.

1 Introduction

2 Each week without a COVID-19 vaccine exacts an unimaginable toll on global public health,
3 economic livelihoods, and political stability [1]. At present, well over 100 COVID-19 vaccine
4 candidates have been identified, at least 23 of which are in clinical trials worldwide [2]. Tradi-
5 tional efficacy trial designs take many months and rely on the persistence of high transmission
6 rates around the trial site. However, several trial design options may expedite the process. Two
7 of the leading options are, first, the use of “human challenge studies,” wherein study participants
8 volunteer to be exposed to the virus instead of having scientists wait for participants to be ex-
9 posed to it in their daily lives (potentially accompanied by additional safety testing in a diverse
10 population) [3]; and, second, the integration of smaller Phase II safety and immunogenicity
11 trials into larger Phase III efficacy trials, wherein more study participants receive a vaccine can-
12 didate before data about its safety and immunogenicity are available from a traditional Phase
13 II trial. These approaches may significantly accelerate COVID-19 vaccine development [3–
14 7]. However, they also create risks to participants: in the case of challenge trials, deliberately
15 exposing participants to the virus; and, in the case of integrated trials, exposing additional par-
16 ticipants to a vaccine before Phase II safety and immunogenicity studies have been completed.

17 While technical questions also surround the use of these and related approaches in COVID-
18 19 vaccine testing [8–11], several of the roadblocks to their implementation depend in part on
19 the answer to an empirical question: would members of the general public see these designs as
20 ethical [12–16]?

21 To understand how members of the public informed of the ethical trade-offs involved would
22 view these questions, we measured public opinion on the ethics of accelerated vaccine trial
23 designs.

24 There are at least three reasons why public opinion towards the ethics of these designs is

25 relevant to vaccine trialists, ethicists, and policymakers.

26 First, there has been considerable debate about the ethics of accelerated designs for COVID-
27 19 vaccine trials, particularly human challenge trials [3, 8, 13–18]. Ethicists have held that part
28 of what determines whether proposed research is ethical are societal attitudes toward that re-
29 search, especially among those individuals and communities who will be most affected by it
30 [19]. This is one reason that community members have long been included on institutional re-
31 search ethics committees [20]. Decisions about whether to use standard or accelerated COVID-
32 19 vaccine trial designs affect people worldwide. Accordingly, the WHO recently identified
33 “consultation and engagement with the public” as a core criterion for the ethical acceptability
34 of challenge trials [12]. “Community engagement” is said to be an important ethical safeguard
35 “to . . . ensure that the research is consistent with the community’s values, show respect for
36 members of the community, and enhance transparency” [19]. Notable scholars have also ex-
37 plicitly called for “public opinion surveys [to] identify concerns” on the ethics of COVID-19
38 challenge trials [13] and “public engagement” to “assess the local acceptability of human chal-
39 lenge studies” [16]. Our research answers these calls.

40 Second, vaccines are only effective if the public gets vaccinated, and public trust in vaccines
41 may depend in part on public views about the ethics of the process by which those vaccines
42 were developed and tested [21]. As a result, some scholars have expressed concern that the
43 use of human challenge trials could “feed distrust among the public,” and therefore “exacerbate
44 challenges in vaccine roll-out and delay uptake of an effective vaccine” [22]. Our research can
45 help empirically assess such concerns.

46 Third, as many decisions about vaccine trials and vaccine manufacturing will be made by
47 government officials who are accountable to public opinion, it is important to know what the
48 public thinks about the social and ethical trade-offs these options involve [23]. In practice, pol-
49 icymakers and regulators may be hesitant to support an accelerated design viewed as unethical

50 by the public.

51 However, despite a plethora of research on public opinion regarding medical research [23]
52 and on how individuals make ethical judgments [24], relatively little is known about how in-
53 dividuals perceive the ethics of accelerated vaccine trials on consenting volunteers, especially
54 in the setting of the current global public health emergency. One exception is Gbesemete et
55 al., who find strong public support for COVID-19 challenge trials in focus groups conducted
56 with 57 young people in the United Kingdom [25]. Our work, described below, builds on their
57 work with a larger sample size, cross-national data, a more diverse subject pool, and a different,
58 complementary research design.

59 **Materials and Methods**

60 To measure public opinion about COVID-19 vaccine testing, we surveyed people worldwide
61 in May 2020. We recruited participants to the survey using the online sample provider Lucid.
62 Prior research shows that US Lucid respondents demographically track well with US national
63 benchmarks and that many political, psychological, and experimental results replicate on Lucid
64 samples [26]. In addition to surveying $n = 2,180$ individuals in the United States, we also
65 surveyed approximately 500 individuals through Lucid in each of the following predominantly
66 English-speaking geographies worldwide: Australia ($n = 500$), Canada ($n = 687$), Hong Kong
67 ($n = 422$), New Zealand ($n = 498$), South Africa ($n = 548$), Singapore ($n = 520$), the United
68 Kingdom ($n = 565$). Survey respondents in Australia, the United Kingdom, and the United
69 States surveys were selected to match Census population benchmarks on age, gender, and race
70 and ethnicity, but this was not possible in the other geographies. Although the results are not
71 meaningfully different across geographies, respondents in Canada, Hong Kong, New Zealand,
72 South Africa, and Singapore were not selected to match Census benchmarks and the results
73 for these geographies should be viewed with this limitation in mind. Furthermore, prior work

74 on the representativeness of Lucid has been limited to the US. All surveys were conducted in
75 English. The Supplementary Materials provides further detail on the demographics of survey
76 respondents. Later, we also report results from sensitivity analyses and placebo tests to further
77 support the robustness of our findings.

78 Survey respondents were randomly assigned to participate in one of two studies. Study 1
79 asked respondents about a standard Phase III trial in which scientists wait for participants to be
80 exposed to the virus in the real world and a challenge trial in which participants are intentionally
81 exposed to the virus. Study 2 asked respondents about a standard Phase II followed by Phase III
82 trial and an integrated Phase II and III trial in which the Phase II safety and immunogenicity trial
83 is integrated into a larger Phase III efficacy trial, thereby reducing the vaccine testing timeline.¹

84 In both studies, we described essential details of these research designs for respondents—
85 including their procedures, risks to participants, and benefits to society. Figure 1 provides an
86 example description of the designs that we showed participants in Study 1 (see Figure S4 for the
87 stimulus used for the integrated trial). We followed these descriptions with a bulleted summary
88 of the key differences (given in the Supplementary Materials).

89 In both studies, we randomly assigned each respondent to several parameters of the vac-
90 cine trials we described, including their sample sizes, COVID-19 infection and death rates for
91 their participants, the vaccine approval timeline under standard trials, and how much alternative
92 designs would accelerate this timeline. These randomizations were intended to evaluate the
93 robustness of our findings to a range of assumptions about the likely design and consequences
94 of vaccine trials given the uncertainty surrounding all of these parameters. The Supplementary
95 Materials provides the full wording of the scenarios.

¹We asked respondents about an integrated design that involved giving additional participants a potentially ineffective or unsafe vaccine before its lack of effectiveness or safety was detected. However, adaptive trial designs exist that would likely result in even fewer participants being exposed than in standard designs, including designs that adaptively learn optimal dosage [6]. Based on our results, it seems unlikely that public opinion would be less friendly to a trial design that was both accelerated and involved lower risk to study participants.

96 There is ongoing scientific uncertainty and differences of opinion regarding the likely de-
97 sign, benefits, and risks of both conventional and accelerated COVID-19 vaccine trial designs.
98 This ongoing uncertainty merits several comments about our summaries of the designs for re-
99 spondents. First, our summaries should not be interpreted as making any new scientific claim
100 about the likely trial designs; rather, they represented our best assessment of the available ev-
101 idence as of when the survey was conducted (May 2020). Second, reflecting the scientific
102 uncertainty about likely trial designs, as described above we also randomized a number of as-
103 pects of the scenarios (as shown in Figure 1 and detailed in the Supplementary Materials). As
104 we show below, our results are robust across these potential trial design details. Third we did
105 not include any scenarios in which accelerated designs took as long or longer than conventional
106 designs, as the ethical case for accelerated designs is mainly relevant if they in fact do accelerate
107 vaccine approval timelines (which is a separate question). Finally, our challenge trial scenarios
108 always described it as “very unlikely any study participants would die of coronavirus” based on
109 the available research about the infection fatality rate for young, healthy people, who we state
110 would form the subject pool for a challenge trial [12, 16, 27, 28].²

111 After presenting the study designs, we asked respondents which of the two trials they would
112 prefer to see scientists conduct. We next asked respondents to rate how ethical and how scien-
113 tifically valid they considered each trial design and how likely they would be to take the vaccine
114 if it had been tested using each design, if the vaccine were to be approved. We also asked
115 several questions to measure respondents’ successful comprehension of the study designs; this
116 allows us to determine whether our survey instrument successfully communicated the intended
117 differences between the vaccine trial designs. We ended the survey with several demographic
118 questions, which allow us to identify and separately analyze data for numerous sub-populations,

²In the largest challenge trial shown to participants ($N = 200$), under a high-end estimate of the infection fatality rate for young people in the available research of 0.03%, the probability that zero participants of 100 in the placebo condition infected with SARS-CoV-2 would die is $(1 - 0.0003)^{100} = 97.0\%$.

Figure 1: Example Stimulus, Study 1 (Study A is the Challenge Trial in this Example)

Name	Study A	Study B	Key Differences
Study procedures	<ol style="list-style-type: none"> Recruit 80 healthy volunteers aged 18-30 to participate. Give 40 people the vaccine and another 40 people a placebo. Scientists expose all 80 participants to the coronavirus. Scientists keep these participants in a medical research center where they cannot infect others, are closely monitored, and provided any necessary medical care. Scientists monitor participants for several weeks to learn if the vaccine works. 	<ol style="list-style-type: none"> Recruit 9,000 healthy volunteers to participate. Give 4,500 people the vaccine and another 4,500 people a placebo. All participants then go about their daily lives. Scientists monitor participants for several months to learn if the vaccine works. 	<ul style="list-style-type: none"> In Study B, people catch the coronavirus on their own as they go about their daily lives. In Study A, participants are intentionally exposed to the virus.
Risks to study participants	<ul style="list-style-type: none"> 40 participants in the placebo group are intentionally exposed to coronavirus while quarantined in a medical research center. Because all participants in this study are young people, it is unlikely any would develop serious complications. In addition, participants would likely develop immunity, so are unlikely to get coronavirus again or pass it along to others. 40 participants who received the vaccine are also intentionally exposed to coronavirus. If the vaccine in the study does not work, they would also get coronavirus. In the worst-case scenario, they could get an especially bad case of it. It is very unlikely any study participants would die of coronavirus. 	<ul style="list-style-type: none"> 90 participants in the placebo group are expected to get coronavirus from their daily lives. It is likely 1 participant would die of coronavirus. If the vaccine in the study does not work, 90 participants who received the vaccine would also get coronavirus. In the worst-case scenario, they could get an especially bad case of it. 	<p>If the vaccine works:</p> <ul style="list-style-type: none"> 50 more volunteers catch the coronavirus in Study B. 1 more volunteer would die in Study B.
Expected time until vaccine ready	If the vaccine works, it would start being distributed widely in 6 months from today, November 2020.	If the vaccine works, it would start being distributed widely in 12 months from today, May 2021.	<ul style="list-style-type: none"> If the vaccine works, Study A would allow it to be ready 6 months sooner than Study B.
Benefits to society, if the vaccine works	<p>During the 6 months between now and when the vaccine is ready, it is estimated that:</p> <ul style="list-style-type: none"> 1,200,000 people will die from coronavirus. Millions of people will remain out of work. <p>However, once the vaccine is ready and starts reaching all those who need it in November 2020:</p> <ul style="list-style-type: none"> Very few people will die of coronavirus any more. Daily life and the economy will return to normal. 	<p>During the 12 months between now and when the vaccine is ready, it is estimated that:</p> <ul style="list-style-type: none"> 2,400,000 people will die from coronavirus. Millions of people will remain out of work. <p>However, once the vaccine is ready and starts reaching all those who need it in May 2021:</p> <ul style="list-style-type: none"> Very few people will die of coronavirus any more. Daily life and the economy will return to normal. 	<ul style="list-style-type: none"> With Study A, 1,200,000 lives are saved. With Study A, people can return to work and daily life can return to normal 6 months sooner.

Note: In order to communicate the details and differences between standard and challenge trials, participants in Study 1 were shown a table like the one above. The highlighted elements were randomized across possible values detailed in the Supplementary Materials. The highlights did not appear for respondents. We randomized these parameters given uncertainty about how particular vaccine trials might be conducted, to ensure our findings were not sensitive to any of these parameters. We did not allow participants to move on from the page describing the trial design until at least 60 seconds had gone by. See the Supplementary Materials for an example from Study 2. As we show in the Supplementary Materials (Tables S2 and S6), the vast majority of survey respondents were able to correctly comprehend the key differences between the designs.

119 including vulnerable sub-populations. We also use these demographic questions to construct
120 weights for the US sample, allowing us to test whether our conclusions change when weighting
121 our sample to the demographics of the US population. The Supplementary Materials provides
122 the full question wording.

123 We pre-registered a pre-analysis plan, provided in the Supplementary Materials, that de-
124 tails the pre-specified analyses we planned to conduct, including which subgroups we would
125 examine. The Supplementary Materials also details two minor deviations from our pre-analysis
126 plan.

127 The survey was approved by the Committee for the Protection of Human Subjects at the
128 University of California, Berkeley (#2020-04-13250) and Yale University (#20000281000). In
129 the survey, we first asked for informed consent to participate. 471 participants did not consent
130 and were removed from the survey.

131 All of the individual participant data collected during the study, after de-identification, study
132 protocol, pre-analysis plan, informed consent form, and analytic code will be available imme-
133 diately following publication with no end date to anyone who wishes to access the data for any
134 purpose. Data will be made available indefinitely at <https://osf.io/bgxe4/>.

135 All authors declare no competing interests.

136 **Results**

137 First, we find that a majority of respondents successfully understood the studies we described,
138 as most correctly answered each of several scenario comprehension questions. For example, in
139 Study 1, 84% (95% CI: 82-85%) of respondents correctly stated that the challenge trial involves
140 intentionally infecting study participants with the virus. Similarly, 75% (95% CI: 73-76%) of
141 respondents in Study 2 correctly stated that the standard trial involves additional safety testing
142 not present in the integrated design. Results for additional scenario comprehension questions

143 are presented in the Tables S2 and S6.

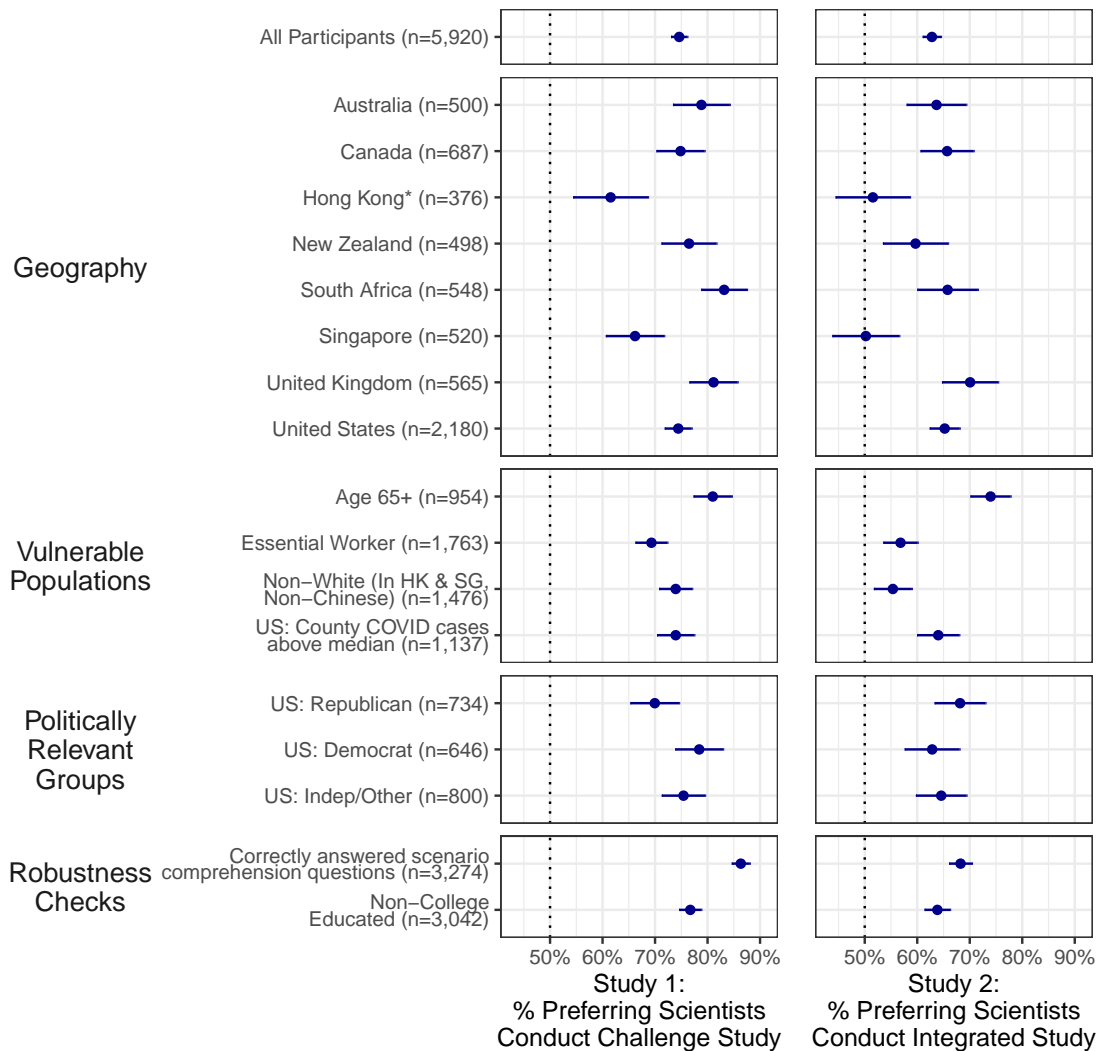
144 We pre-specified that our primary outcome of interest was participants' answer to the ques-
145 tion "If you had to choose, which study would you rather have scientists conduct?" Respondents
146 had the choice of selecting "Study A" or "Study B." Whether the accelerated design was labeled
147 as Study A or B was randomly assigned for each respondent.

148 Figure 2 shows our main results for this primary outcome. Overall, we find broad cross-
149 national support for both the challenge trial and the integrated trial over standard vaccine trials.
150 There are also similar, high levels of support in every subgroup we examined, including among
151 vulnerable populations (e.g., those over 65, essential workers, racial minorities), politically
152 relevant subgroups in the United States, those without a college degree, and among those who
153 correctly answered all scenario comprehension questions asked after the scenario was presented
154 (see also Tables S4 and S8).

155 These results are robust across the various trial design assumptions we randomized to each
156 respondent (e.g., the sample size of each trial we described). Further evidence is presented
157 in the Supplementary Materials. Of particular interest is that the results are not particularly
158 sensitive to the amount of time by which we tell respondents the design could accelerate vaccine
159 development: we find similar results across the range of 2-6 months (see Tables S5 and S9). The
160 results are also robust when the sample from the United States is weighted to match Census
161 targets for gender, race, age, and education (see Tables S3 and S7). Supporting the breadth of
162 public support for accelerating trial designs, multivariate regressions reveal few substantively
163 or statistically significant demographic or attitudinal predictors of support (see Figures S2 and
164 S4).

165 We present further results in the Supplementary Materials for our three secondary outcomes.
166 In Study 1, participants stated that they saw the challenge trial as slightly *more* likely to be
167 ethical ($p < 0.01$), scientifically valid ($p < 0.001$), and that they would be more likely to

Figure 2: Broad Support for Challenge Trials and Integrated Trials That Accelerate COVID-19 Vaccine Development



Notes: The Figure shows the percent of respondents who preferred that scientists conduct each accelerated trial instead of a standard trial. Study 1 considered the use of a challenge trial in which participants are intentionally exposed to the virus instead of, in a standard trial, waiting for participants to be exposed to it in their daily lives. Study 2 considered the integration of smaller Phase II safety and immunogenicity trials into larger Phase III efficacy trials instead of, in a standard trial, waiting for the completion of a Phase II safety and immunogenicity trial before commencing a Phase III efficacy trial. The Figure shows the mean proportion of respondents who say they would prefer scientists to conduct each accelerated design for each study overall, by geography, and across various demographic subgroups. 95% confidence intervals surround the point estimates. Sample sizes shown are totals across both studies; respondents are approximately evenly split across the two studies. See Tables S3, S4, S7, and S8 for numerical values.

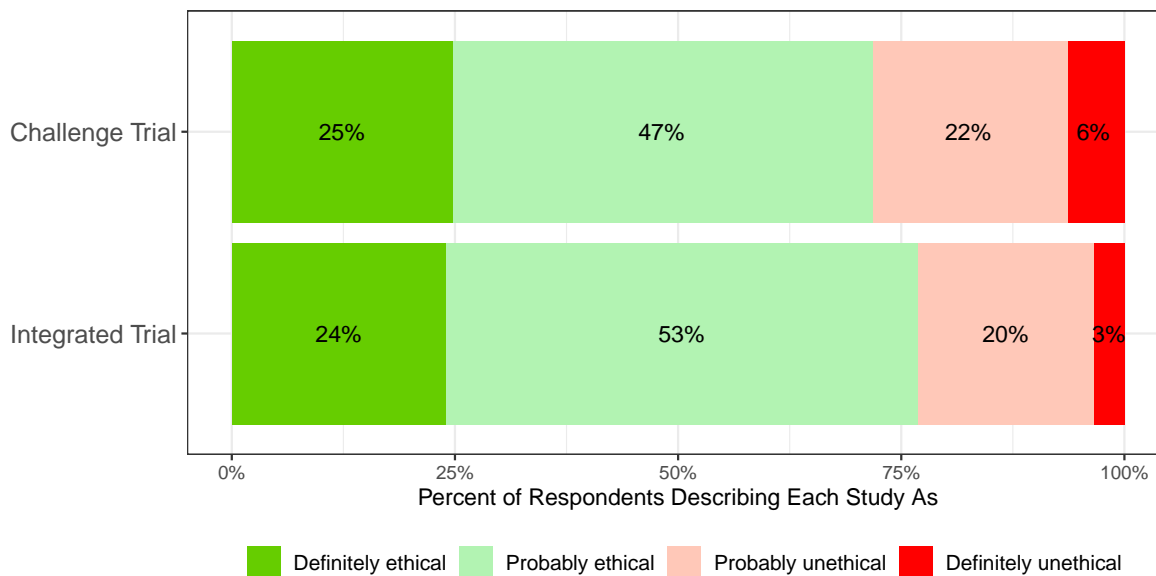
* This Figure presents non-white respondents from Hong Kong only. There were an unanticipatedly large number of participants in Hong Kong who indicated they were white. The 2016 Hong Kong Census estimates that only 0.8% of the Hong Kong population identifies as white; we discuss this issue in further detail in the Supplementary Materials.

168 take an approved vaccine if it had been tested using this design³ ($p < 0.001$). For example,
169 Figure 3 shows the proportions of respondents who described each accelerated design as ethical.
170 72% (95% CI: 70-73%) of participants described the challenge trial as either “probably” or
171 “definitely” ethical, as compared to 68% (95% CI: 67-70%) of participants who described the
172 standard trial as such. These results are broadly consistent across geographies (see Table S3).
173 In Study 2, for the integrated trial, although as described above participants on average are
174 more likely to prefer that scientists conduct the integrated trial instead of the standard trial,
175 the integrated trial has slightly lower averages than the standard trial for our three secondary
176 outcomes (see Table S7). This is because those who prefer the integrated trial rate both trials
177 approximately equally on average, whereas the minority who do prefer the standard trial rate
178 the integrated trial slightly lower on average; however, most in this minority still describe the
179 integrated trial as ethical (58%, 95% CI: 56-61%).

180 To help assess the potential concern that our conclusions spuriously arise from issues with
181 the representativeness of our sample, we also conducted a Rosenbaum sensitivity analysis [29].
182 We find that if supporters and opponents of challenge trials were equally prevalent in the pop-
183 ulation (the null hypothesis), our results finding broader support for challenge trials could only
184 be accounted for by sampling bias if pro-challenge trial individuals were $\gamma = 2.78$ times more
185 likely to be selected for our survey than anti-challenge trial individuals ($\gamma = 1.68$ times more
186 likely for the integrated study). As we discuss in the Supplementary Materials, this is an unusu-
187 ally high γ value for social science studies. Furthermore, the lack of heterogeneity in support
188 for accelerated designs within our sample on observed demographics (e.g., country, age, race,
189 education) provides additional evidence that our results are unlikely to be driven by sampling
190 bias: since none of the many subpopulations we examined in our sample oppose accelerated

³We caution against any strong interpretations of how this result would map to likely behavior, but find it reassuring that participants did *not* on average say they were *less* likely to take the vaccine if a challenge trial were used. Moreover, the minority of the sample that said they thought vaccines were “dangerous” in general were essentially indifferent between the trial types.

Figure 3: Vast Majority of Survey Respondents Describe Both Accelerated Designs as Ethical



191 designs on average, this suggests it is unlikely that there exists a subpopulation very strongly
 192 opposed to accelerated designs that we undersample by a factor of 2.78 or 1.68 (for challenge
 193 and integrated, respectively).

194 To further assess this potential concern, we also conducted a placebo test by comparing our
 195 US sample with nationally-representative samples from the Pew Research Center and Gallup
 196 on items related to vaccine safety and scientific knowledge. First, Gallup data finds that 86%
 197 of Americans say that vaccines are not “more dangerous than the diseases they are designed to
 198 prevent” [30]; in our US sample, this number is 83%. Second, probing scientific knowledge
 199 more generally, Pew data finds that 76% of Americans know that ocean tides are caused by the
 200 pull of the moon and that 72% of Americans know that cell phones use radio waves [31]; in our
 201 US sample, these figures are 77% and 73%, respectively. That our sample closely reflects other,
 202 nationally representative samples on these benchmarks is encouraging for the external validity
 203 of our findings.

204 The results of the sensitivity analysis, which bounds unobserved bias, the lack of observed
205 heterogeneity of responses across observable characteristics, and the close resemblance of our
206 sample to other samples on related dimensions therefore all indicate that sampling bias is very
207 unlikely to account for our results.

208 **Discussion**

209 We conducted large-N surveys in eight countries to gather systematic data about public support
210 for accelerated COVID-19 vaccine trial designs. Our results suggest that there is broad pub-
211 lic support across nations and demographics for accelerating COVID-19 vaccine trials through
212 integrated Phase II and III trials, and especially through challenge trials. A majority of our
213 respondents preferred the use of these accelerated designs and saw them as ethical. Indeed,
214 on average our respondents found challenge trials to be *more* likely to be ethical than standard
215 trials. These results are consistent among every subgroup we examined, including many vul-
216 nerable populations of interest, across the countries we surveyed, and across many plausible
217 vaccine trial design details. While the vast majority of participants in our survey understood
218 the key differences between the designs we described, it is of particular note that, among those
219 who did understand them, fully 86% and 68% preferred for scientists to conduct challenge and
220 integrated trials, respectively.

221 While our results do not at all settle the complex ethical debate about these accelerated trials,
222 nor do they address the scientific or technical questions surrounding their use, they do bolster
223 the case for these trials.

224 First, many scholars and the WHO have called for consulting the public regarding COVID-
225 19 vaccine trial designs [12, 13, 16] given the longstanding view that part of what determines
226 whether proposed research is ethical is whether the affected community supports the research
227 [19, 20]. Our results suggest the public in multiple countries is broadly supportive of accelerated

228 trial designs in this instance and sees them as ethical.

229 Second, our results help diminish the worry that, by appearing unethical, these designs
230 would undermine public trust in vaccines, or in a resulting COVID-19 vaccine. Indeed, respon-
231 dents said they would be *more* likely to take a vaccine that had been tested using a challenge
232 instead of a conventional trial. Although this survey-based self-report should not be interpreted
233 as likely diagnostic of future behavior, it can help assuage concerns that a challenge trial would
234 undermine public willingness to take a COVID-19 vaccine. While our results do not directly
235 speak to how the public would react to a severe adverse event in a challenge trial, it should be
236 noted that such events are also possible in conventional trials, and potentially even likelier to
237 occur given their larger sample sizes.

238 Finally, our results can provide some reassurance to policymakers and regulators concerned
239 about potential negative public sentiment towards the use of these designs.

240 Our results have several limitations worth noting. First, our sample came from an online
241 survey panel. Unfortunately, we determined that gathering a random probability sample would
242 have required significant delay, past the point at which findings about public opinion could
243 inform ongoing decision-making. That said, previous work has found that the sample we used is
244 demographically representative and that a number of political, psychological, and experimental
245 results replicate on it [26]. Moreover, the robustness of our results across geographies and
246 demographic subgroups suggest it is unlikely that a different survey sampling approach in these
247 geographies would produce qualitatively different results [32]. Finally, our sensitivity analyses
248 and placebo tests empirically suggested that sampling bias was very unlikely to account for our
249 results. Still, our data is likely to exclude those without internet access who therefore could not
250 take our survey.

251 Second, although our survey found relatively consistent results in English-speaking coun-
252 tries worldwide, it is possible that results could differ in other locales [24]. The results could

253 also change if they were re-run at a different time. The surveys were fielded in May 2020, prior
254 to the commencement of Phase 3 trials, which may change how the public weighs the risks and
255 benefits of the accelerated designs.

256 Third, when individuals consider moral dilemmas, they typically weigh ethical costs and
257 benefits more, and their initial emotional reactions less, when they engage in more careful con-
258 siderations of the alternatives at hand [33]. Our study participants spent a median of 176 seconds
259 reading about the trial design and answering our primary and secondary outcome questions.
260 This probably reflects less time and engagement than would be present were our respondents to
261 deliberate with others [34], but potentially more than, e.g., when reading a news article. Consis-
262 tent with the possibility that individuals would weigh social benefits even more (and thus prefer
263 accelerated designs to an even greater extent) were they to engage in more deliberate reflection,
264 Gbesemete et al. find even stronger public support for challenge trials in focus groups conducted
265 with 57 young people in the UK [25]. This suggests that methods that facilitate additional par-
266 ticipant education and deliberation, such as focus groups, may find even greater public support
267 for accelerated designs than we found. Future work should nevertheless continue to assess the
268 extent to which conclusions might change if individuals were to encounter information about
269 these trial designs in other settings.

270 Fourth, explaining these trial designs to laypeople necessarily involved making judgments
271 about how to describe the trials amidst scientific uncertainty surrounding the likely designs and
272 their risks. However, we found that our respondents did largely understand the key trade-offs
273 involved and that their support was consistent across the alternative assumptions and scenarios
274 we presented. Nevertheless, scientific and technical knowledge about COVID-19 and associated
275 vaccines is also evolving rapidly, and, although our results were consistent across many design
276 details, it will be critical for future work to explore the robustness of public support for challenge
277 studies and other trial designs across different conditions, especially if assessments of overall

278 risks to participants or benefits to society change dramatically.

279 Finally, although we did demonstrate the robustness of our results across many possible trial
280 design details, there are remaining questions we did not ask respondents and that future research
281 should consider.

282 First, might respondents have preferred the accelerated designs to the standard designs even
283 if they were not faster than the standard designs? For example, in challenge trials, fewer par-
284 ticipants receive the experimental vaccine; participants may prefer the challenge trial for this
285 reason. In our studies, we always presented respondents with scenarios that stated that the ac-
286 celerated designs were expected to produce a vaccine faster than a standard design, so our data
287 do not speak to the question of whether respondents would still prefer the accelerated designs
288 were they not to accelerate the vaccine approval timeline. However, we did find that respon-
289 dents were not particularly sensitive to how many months the accelerated design was able to
290 accelerate the vaccine approval timeline, especially in the case of challenge trials (see Tables
291 S5 and S9). This is consistent with the plausibility of the hypothesis that some respondents pre-
292 ferred the accelerated designs for reasons unrelated to their accelerated timeline. Our research
293 offers a blueprint for future research that can and should consider whether the public finds ac-
294 celerated designs preferable even if they do not accelerate vaccine testing timelines. There are
295 also alternative trial designs that we did not include in our surveys, such as enrolling very large
296 samples in Phase 3 designs to speed up the testing timeline and adaptive designs, that future
297 research could consider as well.

298 Additionally, COVID-19 challenge trials in healthy, young volunteers would probably be
299 supplemented by additional safety testing in more diverse populations; to keep the scenarios
300 relatively concise we did not describe this detail to respondents. Our work also leaves open
301 questions about how the public would react to unforeseen adverse events during accelerated
302 vaccine trials and whether such events would have different effects on public trust than adverse

303 events in conventional trials, public perception of combinations of these accelerated designs,
304 preferences between them, and between them and other accelerated designs, such as adaptive
305 integrated trials [6]. However, our findings do show that it is possible to successfully explain
306 relevant ethical trade-offs in trial design to a global public, and that people in many geographies
307 broadly prefer approaches that accelerate COVID-19 vaccine trial timelines.

308 **Contributors**

309 All authors attest they meet the ICMJE criteria for authorship. All authors contributed to con-
310 ceptualization and study design. DB, JK, and AG wrote the first draft of the manuscript. DB
311 and JK contributed to the data collection and analysis. All authors contributed to critical revi-
312 sion of the manuscript for important intellectual content and gave final approval. DB and JS
313 contributed funding.

314 **Conflicts of Interest**

315 All authors declare no competing interests.

316 **Acknowledgements**

317 The data collection was funded by the University of California, Berkeley. NE and MM are sup-
318 ported by the National Institute of Allergy and Infectious Diseases of the National Institutes of
319 Health (#R01 AI114617). The funders did not influence any aspect of our research or decision
320 to submit.

321 **References**

322 [1] James K. Jackson, Martin A. Weiss, Andres B. Schwarzenberg, and Rebecca M. Nel-
323 son. Global economic effects of covid-19. Technical report, 2020.

- 324 [2] World Health Organization. Draft landscape of covid-19 candidate vac-
325 cines. Available at [https://www.who.int/who-documents-detail/
326 draft-landscape-of-covid-19-candidate-vaccines](https://www.who.int/who-documents-detail/draft-landscape-of-covid-19-candidate-vaccines), July 2020.
- 327 [3] Nir Eyal, Marc Lipsitch, and Peter G Smith. Human challenge studies to accelerate coro-
328 navirus vaccine licensure. *The Journal of Infectious Diseases*, 2020.
- 329 [4] Meta Roestenberg, Ingrid Kamerling, and Saco J de Visser. Controlled human infections
330 as a tool to reduce uncertainty in clinical vaccine development. *Frontiers in medicine*, 5:
331 297, 2018.
- 332 [5] Qing Liu and Gordon W Pledger. Phase 2 and 3 combination designs to accelerate drug
333 development. *Journal of the American Statistical Association*, 100(470):493–502, 2005.
- 334 [6] Scott M Berry, Bradley P Carlin, J Jack Lee, and Peter Muller. *Bayesian adaptive methods
335 for clinical trials*. CRC press, 2010.
- 336 [7] Stanley A Plotkin and Arthur Caplan. Extraordinary diseases require extraordinary solu-
337 tions. *Vaccine*, 2020.
- 338 [8] Jon Cohen. Speed coronavirus vaccine testing by deliberately infecting volunteers? not so
339 fast, some scientists warn. *Science*, March 2020.
- 340 [9] Lawrence Corey, John R Mascola, Anthony S Fauci, and Francis S Collins. A strategic
341 approach to covid-19 vaccine r&d. *Science*, 2020.
- 342 [10] Meagan E Deming, Nelson L Michael, Merlin Robb, Myron S Cohen, and Kathleen M
343 Neuzil. Accelerating development of sars-cov-2 vaccines—the role for controlled human
344 infection models. *New England Journal of Medicine*, 2020.

- 345 [11] World Health Organization. Feasibility, potential value and limitations of establishing
346 a closely monitored challenge model of experimental covid-19 infection and illness in
347 healthy young adult volunteers. Technical report, World Health Organization, 2020.
- 348 [12] World Health Organization. Key criteria for the ethical acceptability of covid-19 human
349 challenge studies. Technical report, World Health Organization, 2020.
- 350 [13] Seema K Shah, Franklin G Miller, Thomas C Darton, Devan Duenas, Claudia Emerson,
351 H Fernandez Lynch, Euzebiusz Jamrozik, Nancy S Jecker, Dorcas Kamuya, Melissa Ka-
352 pulu, et al. Ethics of controlled human infection to study covid-19. *Science*, 368(6493):
353 832–834, 2020.
- 354 [14] Adair D Richards. Ethical guidelines for deliberately infecting volunteers with covid-19.
355 *Journal of Medical Ethics*, 2020.
- 356 [15] Ariella Binik. What risks should be permissible in controlled human infection model
357 studies? *Bioethics*, 2020.
- 358 [16] Euzebiusz Jamrozik and Michael J Selgelid. Covid-19 human challenge studies: ethical
359 issues. *The Lancet Infectious Diseases*, 2020.
- 360 [17] Robert Steel, Lara Buchak, and Nir Eyal. Why continuing uncertainties are no reason to
361 postpone challenge trials for coronavirus vaccines. *Journal of Medical Ethics*, 2020.
- 362 [18] Adam J Kolber. Why we (probably) must deliberately infect. *Journal of Law and the*
363 *Biosciences*, *Forthcoming*, 2020.
- 364 [19] SK Shah, J Kimmelman, A Drapkin Lyerly, H Fernandez Lynch, F McCutchan, FG Miller,
365 R Palacios, C Pardo-Villamizar, and C Zorrilla. Ethical considerations for zika virus hu-

- 366 man challenge trials. *National Institutes of Health, National Institute of Allergy and In-*
367 *fectious Diseases: Seattle, WA, USA, 2017.*
- 368 [20] Robert Klitzman. Institutional review board community members: who are they, what do
369 they do, and whom do they represent? *Academic medicine: journal of the Association of*
370 *American Medical Colleges*, 87(7):975, 2012.
- 371 [21] Heidi J Larson, Louis Z Cooper, Juhani Eskola, Samuel L Katz, and Scott Ratzan. Ad-
372 dressing the vaccine confidence gap. *The Lancet*, 378(9790):526–535, 2011.
- 373 [22] Seema K Shah, Franklin G Miller, Thomas C Darton, Devan Duenas, Claudia Emerson,
374 Holly Fernandez Lynch, Euzebiusz Jamrozik, Nancy S Jecker, Dorcas Kamuya, Melissa
375 Kapulu, et al. Unnecessary hesitancy on human vaccine tests-response. *Science (New*
376 *York, NY)*, 369(6500):151, 2020.
- 377 [23] Alan S Gerber, Eric M Patashnik, David Doherty, and Conor Dowling. A national survey
378 reveals public skepticism about research-based treatment guidelines. *Health Affairs*, 29
379 (10):1882–1884, 2010.
- 380 [24] Edmond Awad, Sohan Dsouza, Azim Shariff, Iyad Rahwan, and Jean-François Bonnefon.
381 Universals and variations in moral decisions made in 42 countries by 70,000 participants.
382 *Proceedings of the National Academy of Sciences*, 117(5):2332–2337, 2020.
- 383 [25] D Gbesemete, M Barker, WT Lawrence, D Watson, H de Graaf, and RC Read. Explor-
384 ing the acceptability of controlled human infection with sarscov2—a public consultation.
385 *BMC medicine*, 18(1):1–8, 2020.
- 386 [26] Alexander Coppock and Oliver A McClellan. Validating the demographic, political, psy-
387 chological, and experimental results obtained from a new source of online survey respon-
388 dents. *Research & Politics*, 6(1):2053168018822174, 2019.

- 389 [27] Henrik Salje, Cécile Tran Kiem, Noémie Lefrancq, Noémie Courtejoie, Paolo Bosetti,
390 Juliette Paireau, Alessio Andronico, Nathanaël Hozé, Jehanne Richet, Claire-Lise Dubost,
391 et al. Estimating the burden of sars-cov-2 in france. *Science*, 2020.
- 392 [28] Robert Verity, Lucy C Okell, Ilaria Dorigatti, Peter Winskill, Charles Whittaker, Natsuko
393 Imai, Gina Cuomo-Dannenburg, Hayley Thompson, Patrick GT Walker, Han Fu, et al.
394 Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet*
395 *Infectious Diseases*, 2020.
- 396 [29] Paul R Rosenbaum. *Observational Studies*. Springer, New York, 2002.
- 397 [30] Gallup. Fewer in u.s. continue to see vaccines as impor-
398 tant. Available at [https://news.gallup.com/poll/276929/
399 fewer-continue-vaccines-important.aspx](https://news.gallup.com/poll/276929/fewer-continue-vaccines-important.aspx), 2020.
- 400 [31] Pew Research Center. A look at what the public knows and does not know about sci-
401 ence. Available at [https://www.pewresearch.org/science/2015/09/10/
402 what-the-public-knows-and-does-not-know-about-science/](https://www.pewresearch.org/science/2015/09/10/what-the-public-knows-and-does-not-know-about-science/), 2015.
- 403 [32] Alexander Coppock, Thomas J Leeper, and Kevin J Mullinix. Generalizability of hetero-
404 geneous treatment effect estimates across samples. *Proceedings of the National Academy*
405 *of Sciences*, 115(49):12441–12446, 2018.
- 406 [33] Joshua D Greene, Sylvia A Morelli, Kelly Lowenberg, Leigh E Nystrom, and Jonathan D
407 Cohen. Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*,
408 107(3):1144–1154, 2008.
- 409 [34] Scott YH Kim, Ian F Wall, Aimee Stanczyk, and Raymond De Vries. Assessing the
410 public’s views in research ethics controversies: deliberative democracy and bioethics as
411 natural allies. *Journal of Empirical Research on Human Research Ethics*, 4(4):3–16, 2009.

Supplementary Materials: Broad Cross-National Public Support for Accelerated COVID-19 Vaccine Trials

Contents

Additional Detail on Survey	1
Survey Procedures	1
Demographics	3
Full Survey Questions and Results	3
Study 1: Challenge Trial	3
Study 2: Integrated Trial	12
Pre-Registration	19
Departures from Pre-Registration	19
Pre-Registration Document	19

Additional Detail on Survey

Survey Procedures

Before conducting the surveys, we filed a pre-registration and pre-analysis plan at <https://osf.io/cgbxa/>, the full text of which is available at the end of the SM. We describe two minor departures from our pre-registered pre-analysis plan in the last section.

We then conducted national surveys in Australia, Canada, Hong Kong, New Zealand, South Africa, Singapore, the United Kingdom, and the United States. We recruited participants online using the sample provider Lucid in early May, 2020. Survey respondents in Australia, the United Kingdom, and the United States surveys were selected to match Census benchmarks. We requested 2,000 participants in the United States and 500 for each other geography. Respondents must take the survey on a desktop computer and be able to read English.

In the survey, we first asked for consent to participate. 471 participants did not consent and were removed from the survey. 164 participants failed a first attention check, which asked participants to select “I understand” in response to “For our research, careful attention to survey questions is critical!”; these participants were removed. As a final attention check, we wrote: “People are very busy these days and many do not have time to follow what goes on in the government. We are testing whether people read questions. To show that you’ve read this much, answer both ‘extremely interested’ and ‘very interested.’” We removed 1,636 participants for not reading carefully enough to select both ‘extremely interested’ and ‘very interested’. 895 did not complete the survey. 5,920 participants remained and completed the survey.

We then told participants:

COVID-19, often called coronavirus, has led to hundreds of thousands of deaths and millions of job losses.

Many experts say that people will continue dying of coronavirus and daily life will not return to normal until there is a vaccine. A vaccine would make many people immune to coronavirus, so

they could not catch coronavirus or give it to other people.

Politicians and experts are debating several approaches to developing a coronavirus vaccine. These approaches are all medically and scientifically valid, but have different strengths and weaknesses. Because the coronavirus crisis affects everyone, many politicians and experts want to know what people think about the best approach. We would like to hear your opinion about which of two approaches you think the government and researchers should take. We plan to share our findings, so please take this survey seriously.

Do you understand?

We then gave the following background on vaccines:

To test whether a new coronavirus vaccine is safe and effective, scientists will accept volunteers for a study on vaccines. Scientists then will use a device (like flipping a coin) that randomly assigns volunteers in the study to either receive the vaccine or a placebo, a safe substance like salt water that does not have any vaccine in it. The scientists will then monitor whether study participants who received the vaccine instead of the placebo are less likely to get the coronavirus. Scientists will know the vaccine works if the participants who received the vaccine were less likely to get coronavirus than the people who received the placebo.

Do you understand?

We then asked two questions on people's general views on vaccines: "In your opinion, how important is it that parents get their children vaccinated?" and "Do you think vaccines are more dangerous than the diseases they are designed to prevent, or not?"

We then randomized participants to one of two studies: Study 1, about Challenge Trials, and Study 2, about Integrated Trials. We always began:

We would first like to get your opinion about two hypothetical ways to do a study to determine whether a new coronavirus vaccine is effective.

- There have already been initial studies that gave the vaccine to a small number of volunteers. In these initial studies, the vaccine appeared safe. [*For challenge trial only:* These initial studies also found that the vaccine has a good chance of working, even if it has not been proven effective yet.]
- For both of the studies we will describe, people have already volunteered to participate. Both studies are practical to start right away.
- People who volunteered to participate in the studies were fully informed about what each study involved before they volunteered.
- If any participants get sick during the studies, their medical care and any lost wages would be paid for. They would also have access to remdesivir, a drug that can often, but not always, improve time to recovery.

We described both studies as practical to begin to ensure respondents did not select on the basis of perceived practicability, as the ethical question is only of interest if both trials are practical to conduct.

We show how we described the trial designs to participants and the outcome variables in the context of the studies below. We did not allow participants to move on from the page describing the trial design until at least 60 seconds had gone by.

We finished the survey with a series of demographic questions regarding age, gender, education, political ideology, employment, religiosity, scientific knowledge, race/ethnicity, and, in the United States only, political party identification and zip code.

Demographics

Survey participants reported the demographic characteristics shown in Table S1. As expected for an online sample, the survey participants are slightly more educated than the population at large on average. Later we show the results are consistent for participants without a college degree. Survey respondents in Australia, the United Kingdom, and the United States surveys were selected to match Census benchmarks on age, gender, and race and ethnicity, but this was not possible in the other geographies. Finally, for the US sample, we later present an analysis weighting our sample to be representative of the US population; the findings do not change. Sample sizes and demographics shown are for the entire sample across both studies; respondents were randomly assigned to the two studies with equal probability, so are approximately evenly split across the two studies.

Table S1: Survey Demographics by Geography

	AUS	CAN	HK	NZ	SA	SG	UK	US	All
Average Age	49	47	33	44	36	33	43	49	44
% Female	52	52	48	45	46	47	47	63	54
% College Graduate	43	46	64	39	35	61	47	51	49
% Non-White (In HK & SG, Non-Chinese)	20	27	25	27	53	21	16	21	25
% Employed as an Essential Worker	24	25	67	22	27	43	22	27	30
% Employed as a Non-Essential Worker	25	26	15	28	24	28	31	19	23
% Unemployed due to COVID	5	4	1	4	10	2	3	6	5
% Furloughed due to COVID	4	7	3	8	12	5	12	5	7
N	500	687	422	498	548	520	565	2,180	5,920

For the US sample, relevant questions also closely match national averages from other surveys:

- Gallup data from December 2019 finds that 86% of Americans say that vaccines are not “more dangerous than the diseases they are designed to prevent”; in our US sample, this number is 83%. (In our sample outside the US, this figure is 80%.)
- Pew data from September 2014 finds that 76% of Americans know that ocean tides are caused by the pull of the moon and that 72% of Americans know that cell phones use radio waves; in our US sample, these figures are 77% and 73%, respectively. (In our sample outside the US, these figures are 73% and 75%.)

Full Survey Questions and Results

Study 1: Challenge Trial

Vaccine Study Design Description

After the preamble quoted above, participants were shown a table that looks like the example given in Figure S1. All of the highlighted elements were randomized. We did not allow participants to move on from the page describing the trial design until at least 60 seconds had gone by.

The example in Figure S1 shows the results of one particular randomization. All the highlighted numbers in the example were randomized, as detailed below. (These highlights did not appear for respondents.) In the below, the bolded numbers correspond with the numbers used in the example in Figure S1. We randomized these parameters given uncertainty about how particular vaccine trials might be conducted, to ensure our findings were not sensitive to any of these parameters.

- Standard Design Trial N: (3,000; 5,000; 7,000; **9,000**; 11,000)
 - N in each condition is calculated as half of this number.

Name	Study A	Study B	Key Differences
Study procedures	<ol style="list-style-type: none"> Recruit 80 healthy volunteers aged 18-30 to participate. Give 40 people the vaccine and another 40 people a placebo. Scientists expose all 80 participants to the coronavirus. Scientists keep these participants in a medical research center where they cannot infect others, are closely monitored, and provided any necessary medical care. Scientists monitor participants for several weeks to learn if the vaccine works. 	<ol style="list-style-type: none"> Recruit 9,000 healthy volunteers to participate. Give 4,500 people the vaccine and another 4,500 people a placebo. All participants then go about their daily lives. Scientists monitor participants for several months to learn if the vaccine works. 	<ul style="list-style-type: none"> In Study B, people catch the coronavirus on their own as they go about their daily lives. In Study A, participants are intentionally exposed to the virus.
Risks to study participants	<ul style="list-style-type: none"> 40 participants in the placebo group are intentionally exposed to coronavirus while quarantined in a medical research center. Because all participants in this study are young people, it is unlikely any would develop serious complications. In addition, participants would likely develop immunity, so are unlikely to get coronavirus again or pass it along to others. 40 participants who received the vaccine are also intentionally exposed to coronavirus. If the vaccine in the study does not work, they would also get coronavirus. In the worst-case scenario, they could get an especially bad case of it. It is very unlikely any study participants would die of coronavirus. 	<ul style="list-style-type: none"> 90 participants in the placebo group are expected to get coronavirus from their daily lives. It is likely 1 participant would die of coronavirus. If the vaccine in the study does not work, 90 participants who received the vaccine would also get coronavirus. In the worst-case scenario, they could get an especially bad case of it. 	<p>If the vaccine works:</p> <ul style="list-style-type: none"> 50 more volunteers catch the coronavirus in Study B. 1 more volunteer would die in Study B.
Expected time until vaccine ready	<p>If the vaccine works, it would start being distributed widely in 6 months from today, November 2020.</p>	<p>If the vaccine works, it would start being distributed widely in 12 months from today, May 2021.</p>	<ul style="list-style-type: none"> If the vaccine works, Study A would allow it to be ready 6 months sooner than Study B.
Benefits to society, if the vaccine works	<p>During the 6 months between now and when the vaccine is ready, it is estimated that:</p> <ul style="list-style-type: none"> 1,200,000 people will die from coronavirus. Millions of people will remain out of work. <p>However, once the vaccine is ready and starts reaching all those who need it in November 2020:</p> <ul style="list-style-type: none"> Very few people will die of coronavirus any more. Daily life and the economy will return to normal. 	<p>During the 12 months between now and when the vaccine is ready, it is estimated that:</p> <ul style="list-style-type: none"> 2,400,000 people will die from coronavirus. Millions of people will remain out of work. <p>However, once the vaccine is ready and starts reaching all those who need it in May 2021:</p> <ul style="list-style-type: none"> Very few people will die of coronavirus any more. Daily life and the economy will return to normal. 	<ul style="list-style-type: none"> With Study A, 1,200,000 lives are saved. With Study A, people can return to work and daily life can return to normal 6 months sooner.

Figure S1: Example Stimulus, Study 1 (Study A is Challenge Trial in this example)

- Challenge Trial N (**80**; 100; 200)
 - N in each condition is calculated as half of this number.
- % in Standard Design that are exposed to coronavirus in their daily lives (**2%**, 5%, 20%). This is calculated in the survey flow and not directly shown to participants.
 - The number of participants who catch coronavirus is calculated as the product of this and the size of the standard trial placebo condition.
- % in Standard Design who die of coronavirus if they are exposed (0.5%, **1%**). This is calculated in the survey flow and not directly shown to participants.
 - The number of participants who die of coronavirus is calculated as the product of this and the number of participants who catch coronavirus, described above.
- How long the Standard Design takes to get a vaccine ready: **12** or 18 months.
 - The date (e.g., **May 2021**), is calculated automatically based on the current date. The number of people in society who die (e.g., **2,400,000**) is calculated by multiplying the number of months until a vaccine is ready by 200,000. 200,000 is a fixed variable for both study designs. We selected 200,000 deaths per month as that is approximately the number of COVID-19 deaths in April 2020, and so therefore represents a likely conservative estimate of likely COVID-19 deaths per month in the months ahead.
- How much faster the Challenge Trial is: 2, 4, or **6** months faster.
 - The date the vaccine is ready if a challenge trial is used is the date above minus this number.
- Which design is described as “Study A” or “Study B”. Whichever study was described as “Study A” always was shown in the first column.

We formed these parameters in consultation with experts in vaccine trial design. The “In the worst-case scenario, they could get an especially bad case of it.” language refers to the possibility that “an immune response to a vaccine can predispose an individual to a worse outcome upon infection” [2].

In the “Key Differences” column, the differences are taken based on the randomizations from Studies A and B. We always describe it as possible but unlikely that any participants in the challenge trial would die of coronavirus given findings that the IFR of people ages 18-30 from COVID-19 is 0.03% [6, 7]. (In the largest challenge trial shown to participants (N=200), the probability that zero participants of 100 in the placebo condition would die of coronavirus is $(1 - 0.0003)^{100} = 97\%$.) Table S5, presented later, shows the results are consistent across parameters we used in describing the challenge and standard trials.

After reading this table, we then provided participants with a short summary of the key points. In the example in Figure S1, this would look as follows:

Summary:

- Study **A**:
 - Study participants get coronavirus on their own as they go about their daily lives.
 - **More study participants catch coronavirus and more of them likely die of it.** / More study participants are likely to die of coronavirus. / Fewer study participants catch coronavirus, but more would likely die of it. / *If the numbers are equal between the designs, this bullet is omitted.*
 - It takes longer for the vaccine to be ready, so more people in society generally die of coronavirus.
- Study **B**:
 - Young people volunteer to have scientists expose them to the coronavirus while they are in a medical research center.
 - **Fewer study participants catch coronavirus. It is possible but very unlikely that any study participants would die of it.** / More study participants catch coronavirus. It is possible but very unlikely that any study participants would die of it. / *If the numbers are equal between the designs, this bullet is omitted.*

- The vaccine is ready sooner, so fewer people die of coronavirus in society generally.

Whichever study was randomized to be Study A is always shown first. Which middle bullet in each scenario corresponding to which design has greater or fewer coronavirus cases and deaths in the trial is determined based on the randomization described above. The highlighted sections correspond with the version that would have been shown for the example in Figure S1.

Respondents then answered the questions that constitute our outcome measures. Next, on a separate page, they were asked to answer scenario comprehension questions to ensure they understood the studies. Respondents could not return to the table when answering the scenario comprehension questions. Below we describe the outcome measures and scenario comprehension questions in more detail.

Scenario Comprehension Outcomes

We asked the following scenario comprehension questions to ensure participants, 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?
- 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?

Responses are coded as “1” if the participant provided the correct answer and “0” if they coded the incorrect answer. Table S2 reports the average rate of correct answers overall and by geography.

There were an unanticipatedly large number of participants in Hong Kong who identified as white, and these participants gave distinctive responses to all the questions, affecting our average characterization of Hong Kong. The 2016 Hong Kong Census estimates that only 0.8% of the Hong Kong population identifies as white [1], so we report the results for self-identified non-white and white Hong Kong participants separately. As Tables S2 and S6 show, self-identified whites in Hong Kong were especially unlikely to understand the scenarios correctly, suggesting this group of self-identified white Hong Kong residents may have been a subset of participants in Hong Kong who were answering the survey carelessly, including the racial identification question itself. Consistent with this interpretation, the median self-identified white participant in Hong Kong spent only 95 seconds reading the main study table (we forced participants to spend at least 60 seconds before they could advance), versus a median of 124 seconds in the rest of the Hong Kong sample and a median of 181 in the sample outside of Hong Kong.

Table S2 also shows the results for the US when weighting the sample to the 2019 US Census population estimates for gender, age, race, and education. We construct these weights using entropy balancing [3].

The p -values in Table S2 are from one-sample t -tests testing the null hypothesis that the rate of correct answers is equal to 0.5, which is what would be expected from random guessing.

The vast majority of participants understood the scenarios. Later, in Table S4, we show that support for the challenge trial is strongest for those who correctly answered all the scenario comprehension questions.

Table S2: Study 1 - Proportions Correctly Answering Scenario Comprehension Questions

	Participants Intentionally Infected In Challenge Correct (0/1)			Vaccine Is Ready Faster With Challenge Correct (0/1)			More Participants Die In Which Study Correct (0/1)			More People in Society Die In Which Study Correct (0/1)			N
	Mean	95% CI	p	Mean	95% CI	p	Mean	95% CI	p	Mean	95% CI	p	
All Participants	0.84	0.82 - 0.85	0	0.83	0.81 - 0.84	0	0.67	0.65 - 0.69	0	0.83	0.81 - 0.84	0	2,988
AUS	0.85	0.80 - 0.89	0	0.83	0.78 - 0.88	0	0.66	0.59 - 0.72	0	0.84	0.79 - 0.89	0	222
CAN	0.85	0.81 - 0.89	0	0.85	0.81 - 0.89	0	0.68	0.63 - 0.73	0	0.83	0.79 - 0.87	0	346
HK (Non-White)	0.75	0.68 - 0.81	0	0.71	0.65 - 0.78	0	0.57	0.50 - 0.64	0.05	0.68	0.61 - 0.74	0	182
HK (White)	0.58	0.37 - 0.78	0.44	0.5	0.29 - 0.71	1	0.54	0.33 - 0.74	0.7	0.58	0.37 - 0.78	0.44	26
NZ	0.82	0.78 - 0.87	0	0.82	0.77 - 0.87	0	0.63	0.57 - 0.69	0	0.83	0.79 - 0.88	0	255
SA	0.88	0.85 - 0.92	0	0.89	0.86 - 0.93	0	0.72	0.67 - 0.78	0	0.91	0.87 - 0.94	0	285
SG	0.79	0.75 - 0.84	0	0.78	0.73 - 0.82	0	0.59	0.54 - 0.65	0	0.79	0.74 - 0.83	0	281
UK	0.84	0.80 - 0.88	0	0.88	0.84 - 0.91	0	0.75	0.70 - 0.80	0	0.87	0.83 - 0.91	0	281
US	0.85	0.83 - 0.87	0	0.83	0.81 - 0.85	0	0.69	0.66 - 0.71	0	0.83	0.81 - 0.86	0	1,110
US, Weighted	0.83	0.81 - 0.85	0	0.82	0.79 - 0.84	0	0.66	0.63 - 0.69	0	0.81	0.78 - 0.83	0	1,110

Primary and Secondary Outcomes

We asked the following outcome measures:

- If you had to choose, which study would you rather have scientists conduct?
- How ethical do you think the studies are? (Asked for both trial designs)
- How scientifically valid do you think the studies are? (Asked for both trial designs)
- 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 trial designs)

We pre-specified the first as a primary outcome and the difference between the ratings for the challenge and standard trial designs as secondary outcomes. This difference indicates the extent to which the average respondent judged the challenge trial to be more ethical, scientifically valid, or likelier to be taken than the standard trial, with higher values indicating the challenge trial is more ethical, scientifically valid, or likelier to be taken.

Table S3 shows the overall results and results by geography on each of these dependent variables. We again separately report results for self-identified whites and non-whites in Hong Kong for the reasons described above. The p -value for the primary outcome is from one-sample t -tests testing the null hypothesis that the preference for challenge trials equals the preference for the standard design (0.5), which would indicate an equal number of participants selecting the challenge and standard trial designs. For the other outcomes, the p -values are from one-sample t -tests testing the null hypothesis that the mean is equal to 0, meaning there is no difference between respondents' ratings of the challenge trial and standard design on the outcome measure at hand.

Table S3: Study 1 - Main Results, All Participants and by Participant Geography

	Prefers Scientists Conduct Challenge Study (0/1)			How Ethical (1-4): Challenge Minus Standard			How Scientifically Valid (1-4): Challenge Minus Standard			Would Take Vaccine (1-4): Challenge Minus Standard			N
	Mean	95% CI	p	Mean	95% CI	p	Mean	95% CI	p	Mean	95% CI	p	
All Participants	0.75	0.73 - 0.76	0	0.06	0.02 - 0.10	0.01	0.23	0.20 - 0.27	0	0.14	0.11 - 0.17	0	2,988
AUS	0.79	0.73 - 0.84	0	0.05	-0.10 - 0.19	0.53	0.23	0.12 - 0.34	0	0.07	-0.03 - 0.16	0.15	222
CAN	0.75	0.70 - 0.79	0	-0.05	-0.16 - 0.06	0.39	0.14	0.05 - 0.24	0	0.05	-0.02 - 0.13	0.16	346
HK (Non-White)	0.62	0.54 - 0.69	0	-0.02	-0.20 - 0.17	0.86	0.18	0.03 - 0.32	0.02	0.1	-0.05 - 0.25	0.2	182
HK (White)	0.42	0.22 - 0.63	0.44	0.5	0.08 - 0.92	0.02	-0.27	-0.71 - 0.17	0.22	0	-0.47 - 0.47	1	26
NZ	0.76	0.71 - 0.82	0	0.18	0.04 - 0.31	0.01	0.29	0.18 - 0.40	0	0.25	0.16 - 0.35	0	255
SA	0.83	0.79 - 0.88	0	0.25	0.09 - 0.41	0	0.49	0.37 - 0.61	0	0.37	0.27 - 0.47	0	285
SG	0.66	0.61 - 0.72	0	-0.1	-0.23 - 0.04	0.17	0.17	0.05 - 0.28	0	0.07	-0.04 - 0.17	0.21	281
UK	0.81	0.77 - 0.86	0	0.1	-0.04 - 0.24	0.15	0.23	0.13 - 0.33	0	0.09	-0.00 - 0.18	0.06	281
US	0.74	0.72 - 0.77	0	0.05	-0.02 - 0.11	0.14	0.22	0.16 - 0.28	0	0.15	0.10 - 0.20	0	1,110
US, Weighted	0.73	0.71 - 0.76	0	0.07	0.00 - 0.14	0.04	0.21	0.16 - 0.27	0	0.14	0.09 - 0.19	0	1,110

Table S4 shows that the results are similar across a variety of vulnerable populations, those who answered all the scenario comprehension questions correctly, and various politically relevant subgroups. We also examine a racial subgroup that is defined as non-white in Australia, Canada, New Zealand, the United Kingdom, and the United States, and as non-Chinese in Hong Kong and Singapore. The not college educated subgroup was not pre-specified, but is shown to demonstrate that our conclusions are not driven by the fact that our sample is slightly more educated than the general population.

Table S4: Study 1 - Main Results within Demographic Subgroups

	Prefers Scientists Conduct Challenge Study (0/1)			How Ethical (1-4): Challenge Minus Standard			How Scientifically Valid (1-4): Challenge Minus Standard			Would Take Vaccine (1-4): Challenge Minus Standard			N
	Mean	95% CI	p	Mean	95% CI	p	Mean	95% CI	p	Mean	95% CI	p	
All Participants	0.75	0.73 - 0.76	0	0.06	0.02 - 0.10	0.01	0.23	0.20 - 0.27	0	0.14	0.11 - 0.17	0	2,988
Vulnerable Populations													
Age 65+	0.81	0.77 - 0.85	0	0.17	0.08 - 0.26	0	0.28	0.19 - 0.36	0	0.17	0.10 - 0.25	0	447
Essential Worker	0.69	0.66 - 0.72	0	-0.01	-0.08 - 0.07	0.88	0.19	0.13 - 0.26	0	0.09	0.03 - 0.15	0	883
Non-White (In HK & SG, Non-Chinese)	0.74	0.71 - 0.77	0	0.05	-0.04 - 0.15	0.25	0.23	0.15 - 0.30	0	0.2	0.14 - 0.27	0	748
US: County cases >median	0.74	0.70 - 0.77	0	0.04	-0.05 - 0.14	0.37	0.22	0.14 - 0.30	0	0.17	0.10 - 0.24	0	587
Political Groups (US Only)													
US: Republican	0.7	0.65 - 0.75	0	-0.01	-0.12 - 0.10	0.85	0.18	0.08 - 0.27	0	0.09	0.01 - 0.17	0.02	376
US: Democrat	0.78	0.74 - 0.83	0	0.15	0.03 - 0.27	0.01	0.25	0.15 - 0.36	0	0.21	0.12 - 0.30	0	315
US: Indep/Other	0.75	0.71 - 0.80	0	0.03	-0.09 - 0.14	0.65	0.23	0.14 - 0.33	0	0.16	0.07 - 0.24	0	419
Robustness Checks													
Correctly answered all comprehension questions	0.86	0.85 - 0.88	0	0.17	0.11 - 0.22	0	0.37	0.33 - 0.42	0	0.23	0.19 - 0.26	0	1,529
No college degree	0.77	0.75 - 0.79	0	0.15	0.09 - 0.20	0	0.28	0.23 - 0.33	0	0.18	0.14 - 0.23	0	1,546

Table S5 shows the results are consistent across parameters we used in describing the challenge and standard trials.

Table S5: Study 1 - Main Results, By Randomized Description of Trial Designs

	Prefers Scientists Conduct Challenge Study (0/1)			How Ethical (1-4): Challenge Minus Standard			How Scientifically Valid (1-4): Challenge Minus Standard			Would Take Vaccine (1-4): Challenge Minus Standard			N
	Mean	95% CI	p	Mean	95% CI	p	Mean	95% CI	p	Mean	95% CI	p	
All participants	0.75	0.73 - 0.76	0	0.06	0.02 - 0.10	0.01	0.23	0.20 - 0.27	0	0.14	0.11 - 0.17	0	2,988
Speed of Standard Trial; How Many Months Challenge Accelerates Timeline													
12; Challenge 2 Mo Faster	0.73	0.69 - 0.77	0	0.01	-0.09 - 0.11	0.82	0.23	0.15 - 0.31	0	0.1	0.02 - 0.17	0.01	513
12; Challenge 4 Mo Faster	0.75	0.72 - 0.79	0	0.1	-0.01 - 0.21	0.07	0.27	0.18 - 0.36	0	0.17	0.09 - 0.24	0	477
12; Challenge 6 Mo Faster	0.78	0.74 - 0.81	0	0.07	-0.03 - 0.17	0.19	0.29	0.21 - 0.37	0	0.17	0.10 - 0.25	0	509
18; Challenge 2 Mo Faster	0.72	0.68 - 0.76	0	0.04	-0.06 - 0.14	0.42	0.15	0.06 - 0.24	0	0.13	0.05 - 0.21	0	466
18; Challenge 4 Mo Faster	0.76	0.73 - 0.80	0	0.14	0.04 - 0.24	0.01	0.3	0.22 - 0.38	0	0.13	0.06 - 0.20	0	483
18; Challenge 6 Mo Faster	0.74	0.70 - 0.77	0	-0.01	-0.10 - 0.09	0.88	0.16	0.08 - 0.23	0	0.17	0.10 - 0.23	0	540
% of Placebo Subjects in Standard Trial Who Get Sick, Die													
2% Sick; 0.5% of Sick Die	0.71	0.67 - 0.75	0	-0.16	-0.25 - -0.06	0	0.21	0.13 - 0.29	0	0.06	0.00 - 0.13	0.05	500
2% Sick; 1% of Sick Die	0.75	0.71 - 0.78	0	0.02	-0.08 - 0.12	0.68	0.23	0.15 - 0.31	0	0.17	0.10 - 0.24	0	515
5% Sick; 0.5% of Sick Die	0.75	0.71 - 0.79	0	0.12	0.02 - 0.23	0.02	0.25	0.17 - 0.33	0	0.19	0.12 - 0.26	0	488
5% Sick; 1% of Sick Die	0.76	0.72 - 0.80	0	0.18	0.07 - 0.28	0	0.26	0.18 - 0.35	0	0.16	0.09 - 0.24	0	488
20% Sick; 0.5% of Sick Die	0.73	0.69 - 0.77	0	0.11	0.00 - 0.21	0.04	0.21	0.12 - 0.30	0	0.12	0.04 - 0.20	0.01	476
20% Sick; 1% of Sick Die	0.78	0.74 - 0.82	0	0.08	-0.01 - 0.18	0.09	0.23	0.15 - 0.31	0	0.17	0.09 - 0.24	0	521
Size of Standard Trial													
Standard Trial N = 3000	0.74	0.70 - 0.77	0	0.01	-0.08 - 0.10	0.86	0.27	0.20 - 0.35	0	0.1	0.03 - 0.17	0.01	605
Standard Trial N = 5000	0.77	0.73 - 0.80	0	0.08	-0.02 - 0.17	0.1	0.24	0.17 - 0.31	0	0.16	0.10 - 0.23	0	620
Standard Trial N = 7000	0.75	0.71 - 0.78	0	0.05	-0.04 - 0.14	0.3	0.21	0.14 - 0.29	0	0.16	0.10 - 0.23	0	596
Standard Trial N = 9000	0.75	0.71 - 0.78	0	0.08	-0.01 - 0.18	0.08	0.2	0.12 - 0.28	0	0.14	0.07 - 0.20	0	570
Standard Trial N = 11000	0.73	0.69 - 0.77	0	0.07	-0.02 - 0.17	0.12	0.23	0.16 - 0.31	0	0.16	0.10 - 0.23	0	597

The only case in which participants rate the standard trial as more ethical than the challenge trial is for participants who saw a version of the trial summary where the proportion of subjects in the standard trial who are sick was randomized to the lowest (2%) and the proportion of those who get sick who die was randomized to the lowest (0.5%). Respondents' ratings of the standard trial as more ethical than the challenge trial in this case is driven by participants rating the standard trials as more ethical, not the challenge trial as less ethical, than participants who were randomized to other scenarios. Moreover, it is driven by participants who saw smaller standard trials (Ns of 3,000 or 5,000); in these cases, no participants in the standard trials

die, and fewer participants catch coronavirus than in the challenge trial. It is possible that participants recognized an ethical advantage of the standard trial in this case. However, even these participants still rated the challenge trial similarly ethical as other participants and still favored the challenge trial on average.

Multivariate Regression Results

Figure S2 also visualizes coefficients from a multivariate regression predicting preference for the challenge trial. Variables denoted with (Std.) have been rescaled to standard deviation 1. Note that these coefficients should be interpreted as descriptive, not causal, and that all coefficients are calculated holding constant the other variables shown. The intercept of the regression is 0.78 and the omitted base categories in the regressions are: United States; age 18-24; male; no college degree; employment one of homemaker, student, unable to work, or unemployed since before the coronavirus pandemic; the challenge trial being study A; and average values of the rescaled variables. The coefficients below represent estimated differences from these base categories, holding constant on the other variables in the model.

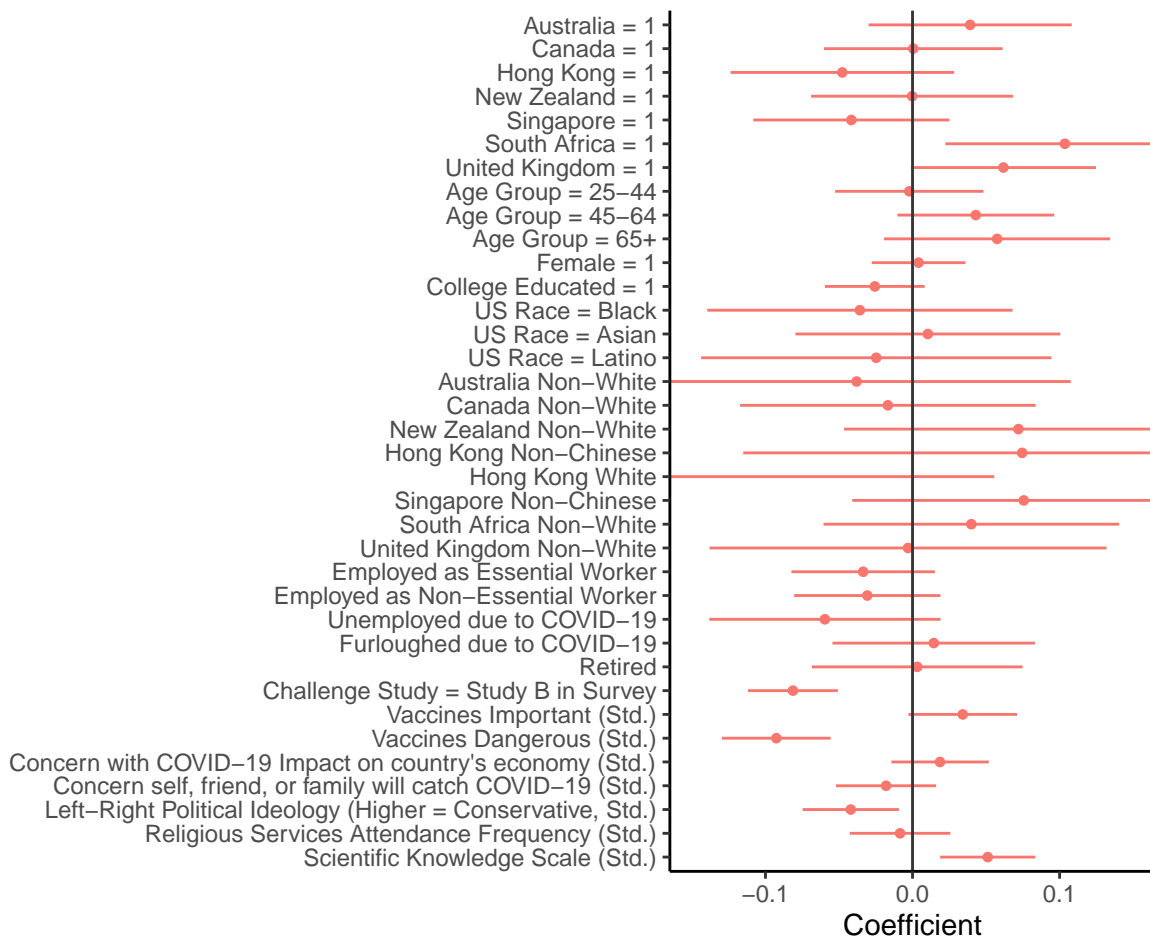


Figure S2: Predictors of Preference for Challenge Trial

Figure S3 shows the results of a multivariate regression regressing the same variables as above on a binary indicator for whether respondents got all the comprehension questions about the challenge study correct. Gender, age, views towards vaccines, and scientific knowledge generally predict comprehension. This regression was not pre-registered.

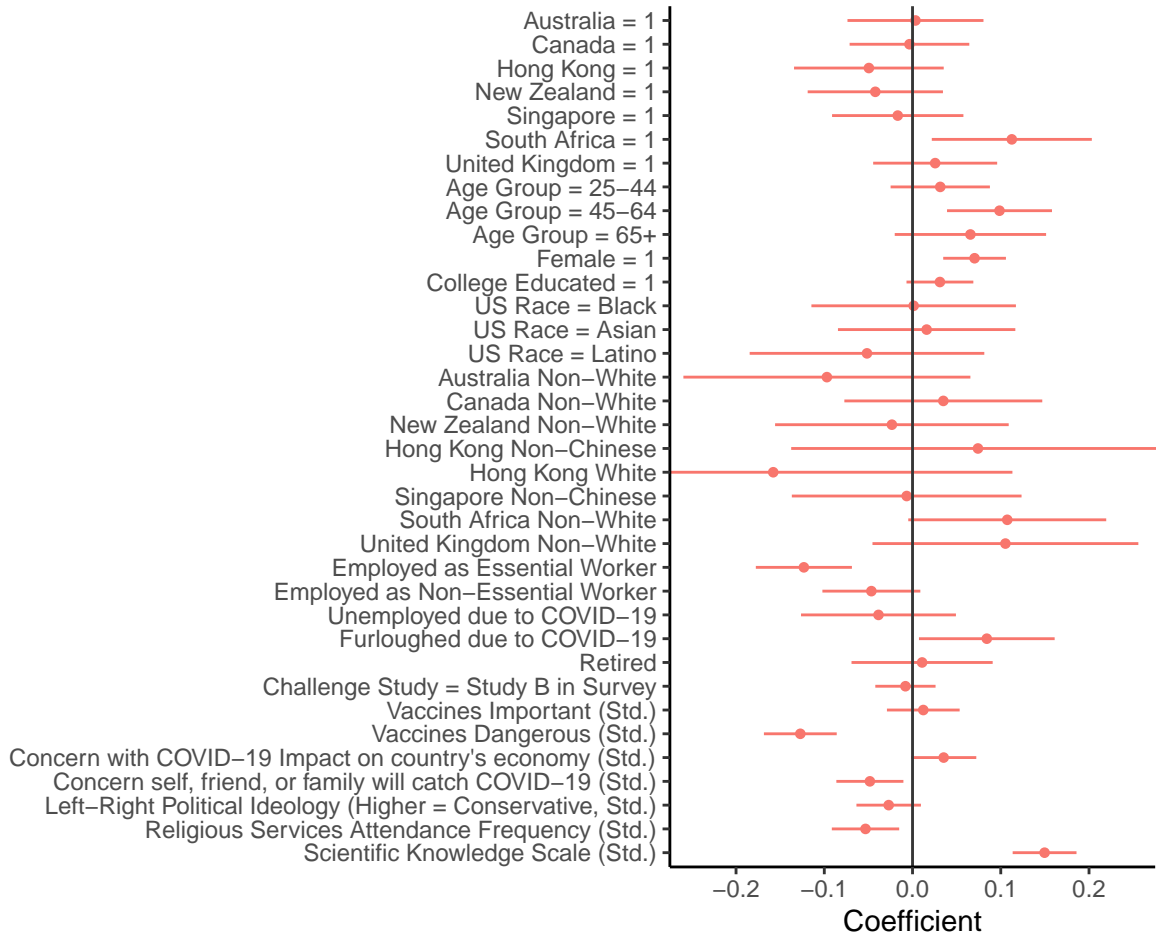


Figure S3: Predictors of Answering Scientific Knowledge Questions Correctly for Challenge Trial

Sensitivity Analysis

We also conduct a sensitivity analysis to determine how unrepresentative our sample would need to be in order to alter our conclusions. In particular, we compute [5]'s γ for the proportion that prefer the challenge study as 2.78 at a p-value of 0.10. This means that our statistical significance for the null hypothesis that challenge trials are equally preferred to standard trials would only no longer reach significance at the 0.10 level were individuals who do not prefer challenge trials to be 2.78 times less likely than individuals who prefer challenge trials to be selected for the survey. For comparison, it is very unusual in social science studies to observe a γ value greater than 2 [4].

Study 2: Integrated Trial

Vaccine Study Design Description

After the preamble quoted earlier, participants were shown a table that looks like the example in Figure S4. All of the highlighted elements were randomized. We did not allow participants to move on from the page describing the trial design until at least 60 seconds had gone by.

Name	Study A	Study B	Key Differences
Study procedures	<ol style="list-style-type: none"> 1. Recruit 9,000 healthy volunteers to participate. 2. Give a few hundred participants the vaccine every week. Stop the study and stop giving participants the vaccine if it is found to be unsafe or unlikely to work. 3. If the vaccine is safe, eventually 4,500 participants will get the vaccine. Another 4,500 will get a placebo. 	<ol style="list-style-type: none"> 1. In a preliminary study, recruit 200 healthy people who volunteer to receive the vaccine. After a month, check whether the vaccine has a good chance of working and gather more data on how safe it is. If the vaccine is found to be unsafe or is shown to be unlikely to work, the study stops here. 2. Recruit 9,000 more healthy volunteers to participate. 3. Give 4,500 participants the vaccine. Another 4,500 get a placebo. 	<ul style="list-style-type: none"> • Study B first takes a few months to determine whether the vaccine has a good chance of working and further checks whether it is safe before continuing. Study A immediately starts investigating whether the vaccine actually works, checking to see whether the vaccine is safe and has a good chance of working as it gets started.
Risks to study participants	<ul style="list-style-type: none"> • While unlikely, if it turns out the vaccine was not safe, approximately 250 people would have received the unsafe vaccine before scientists realized it was unsafe and could have negative side effects. 	<ul style="list-style-type: none"> • If it turns out the vaccine was not safe, 200 people would have received the unsafe vaccine and could have negative side effects. 	<ul style="list-style-type: none"> • While unlikely, if the vaccine is not safe, Study A would expose 50 additional people to it before the problem was found.
Expected time until vaccine ready	<p>If the vaccine works, it would start being distributed widely starting in 10 months, March 2021.</p>	<p>If the vaccine works, it would start being distributed widely in 12 months, May 2021.</p>	<ul style="list-style-type: none"> • If the vaccine works, Study A would allow it to be ready 2 months sooner.
Benefits to society, if the vaccine works	<p>During the 10 months between now and when the vaccine is ready, it is estimated that:</p> <ul style="list-style-type: none"> • 2,000,000 people will die from coronavirus. • Millions of people will remain out of work. <p>However, once the vaccine is ready and starts reaching all those who need it in March 2021:</p> <ul style="list-style-type: none"> • Very few people will die of coronavirus any more. • Daily life and the economy will return to normal. 	<p>During the 12 months between now and when the vaccine is ready, it is estimated that:</p> <ul style="list-style-type: none"> • 2,400,000 people will die from coronavirus. • Millions of people will remain out of work. <p>However, once the vaccine is ready and starts reaching all those who need it in May 2021:</p> <ul style="list-style-type: none"> • Very few people will die of coronavirus any more. • Daily life and the economy will return to normal. 	<ul style="list-style-type: none"> • With Study A, 400,000 lives are saved. • With Study A, people can return to work and daily life can return to normal 2 months sooner.

Figure S4: Example Stimulus, Study 2 (Study A is Integrated Trial in this example)

The example in Figure S4 shows the results of one particular randomization. All the highlighted numbers in the example were randomized, as detailed below. (These highlights did not appear to survey respondents.) In the below, the bolded numbers correspond with the numbers in the example shown in Figure S4. We

randomized these parameters given uncertainty about how particular vaccine trials might be conducted, to ensure our findings were not sensitive to any of these parameters.

The following elements could have been randomized. The bold corresponds to the example shown in Figure S4.

- Standard Design Trial N (3,000; 5,000; 7,000; **9,000**; 11,000)
 - N in each condition is calculated as half of this number.
- How long the Standard Design takes to get a vaccine ready: **12** or 18 months.
 - The date (e.g., **May 2021**), is calculated automatically based on the current date. The number of people in society who die (e.g., **2,400,000**) is calculated by multiplying the number of months until a vaccine is ready by 200,000. 200,000 is a fixed variable for both study designs. We selected 200,000 deaths per month as that is approximately the number of COVID-19 deaths in April 2020, and so therefore represents a likely conservative estimate of likely COVID-19 deaths per month in the months ahead.
- How much faster the Integrated Design is: **2**, 4, or 6 months faster
 - The date the vaccine is ready if an integrated trial is used is the date above minus this number.
- How many people in the Integrated Design are exposed to the vaccine before it stops, in the case that the vaccine is found to be unsafe (**250**; 400; 1,000)
- Which design is described as “Study A” or “Study B”. Whichever study was described as “Study A” always was shown in the first column.

In the “Key Differences” column, the differences are taken based on the randomizations from Studies A and B. Table S9, presented later, shows the results are consistent across parameters we used in describing the integrated and standard trials.

After reading this table, we then provided participants with a short summary of the key points. In the example in Figure S4, this would look as follows:

Summary:

- Study **A**:
 - Scientists first conduct a study with fewer participants to be extra sure the vaccine is safe and likely to work before continuing.
 - It takes longer for the vaccine to be ready, so more people die of coronavirus in society generally.
- Study **B**:
 - Scientists give a larger number of participants the vaccine sooner, collecting data as they go to be extra sure the vaccine is safe and has a good chance of working.
 - There is a small chance that the vaccine is unsafe or won’t work, in which case **250** more people would have received this ineffective or unsafe vaccine.
 - The vaccine is ready sooner, so fewer people die of coronavirus in society generally.

Whichever study was randomized to be Study A is always shown first.

Respondents then answered the primary outcome measures. Next, on a separate page, they were asked to answer scenario comprehension questions to ensure they understood the studies. Respondents could not return to the table when answering the comprehension questions. Below we describe the outcome measures and comprehension questions in more detail.

Scenario Comprehension Outcomes

We asked the following scenario comprehension questions:

- 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?
- Which of the two studies we asked about involves doing additional safety testing on a smaller group first?

Responses are coded as “1” if the participant provided the correct answer and “0” if they coded the incorrect answer. Table S6 reports the average rate of correct answers overall and by geography. We also show the results for the US when weighting the sample to the 2019 US Census population estimates for gender, age, race, and education, as described above. We again separately report results for self-identified whites and non-whites in Hong Kong for the reasons described above.

The p -value is from one-sample t -tests testing the null hypothesis that the rate of correct answers is equal to 0.5, which is what would be expected from random guessing.

The vast majority of participants understood the scenarios. Later, in Table S8, we show that support for the integrated trial is strongest for those who correctly answered all the scenario comprehension.

Table S6: Study 2 - Proportions Correctly Answering Scenario Comprehension Questions

	Vaccine Is Ready Faster With Integrated Correct (0/1)			More People in Society Die In Which Study Correct (0/1)			Standard Involves Additional Testing Correct (0/1)			N
	Mean	95% CI	p	Mean	95% CI	p	Mean	95% CI	p	
All Participants	0.85	0.83 - 0.86	0	0.81	0.79 - 0.82	0	0.75	0.73 - 0.76	0	2,932
AUS	0.84	0.79 - 0.88	0	0.81	0.76 - 0.85	0	0.76	0.71 - 0.81	0	278
CAN	0.86	0.82 - 0.90	0	0.79	0.75 - 0.84	0	0.75	0.71 - 0.80	0	341
HK (Non-White)	0.76	0.70 - 0.82	0	0.73	0.67 - 0.79	0	0.64	0.58 - 0.71	0	194
HK (White)	0.55	0.31 - 0.79	0.67	0.65	0.42 - 0.88	0.19	0.35	0.12 - 0.58	0.19	20
NZ	0.86	0.81 - 0.90	0	0.82	0.77 - 0.87	0	0.78	0.73 - 0.83	0	243
SA	0.93	0.90 - 0.96	0	0.87	0.83 - 0.91	0	0.78	0.73 - 0.83	0	263
SG	0.79	0.74 - 0.84	0	0.8	0.75 - 0.85	0	0.69	0.63 - 0.75	0	239
UK	0.88	0.84 - 0.92	0	0.83	0.79 - 0.88	0	0.79	0.74 - 0.84	0	284
US	0.85	0.83 - 0.87	0	0.79	0.77 - 0.82	0	0.76	0.74 - 0.79	0	1,070
US, Weighted	0.85	0.83 - 0.87	0	0.8	0.78 - 0.83	0	0.76	0.73 - 0.78	0	1,070

Primary and Secondary Outcomes

We asked the following outcome measures:

- If you had to choose, which study would you rather have scientists conduct?
- How ethical do you think the studies are?
- How scientifically valid do you think the studies are?
- 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?

We pre-specified the first as a primary outcome and the difference between the ratings for the integrated and standard trial designs as secondary outcomes.

Table S7 shows the overall results and results by geography on each of these dependent variables. We again separately report results for self-identified whites and non-whites in Hong Kong for the reasons described above. The p -value for the primary outcome is from one-sample t -tests testing the null hypothesis that an equal number of participants selecting the integrated and standard trial designs (0.5). For the other outcomes, the p -values are from one-sample t -tests testing the null hypothesis that the mean is equal to 0, meaning

there is no difference respondents' ratings of the integrated trial and standard design on the outcome measure at hand.

Even though participants on average are more likely to select the integrated trial, the averages of the secondary outcomes are slightly higher for the standard design because those who select the standard design gave more extreme responses. For example, participants who selected the integrated trial rated the integrated trial as equally ethical (only 0.01 scale points less ethical on average), whereas those who favored the standard trial rated the integrated trial as 0.62 scale points less ethical on average. However, 58% of respondents who did not prefer the integrated trial still said they thought the integrated trial was "probably" or "definitely ethical."

Table S7: Study 2 - Main Results, All Participants and by Participant Geography

	Prefers Scientists Conduct Integrated Study (0/1)			How Ethical (1-4): Integrated Minus Standard			How Scientifically Valid (1-4): Integrated Minus Standard			Would Take Vaccine (1-4): Integrated Minus Standard			N
	Mean	95% CI	p	Mean	95% CI	p	Mean	95% CI	p	Mean	95% CI	p	
All Participants	0.63	0.61 - 0.65	0	-0.24	-0.27 - -0.21	0	-0.05	-0.07 - -0.03	0	-0.05	-0.08 - -0.03	0	2,932
AUS	0.64	0.58 - 0.69	0	-0.27	-0.35 - -0.19	0	-0.05	-0.12 - 0.02	0.18	-0.06	-0.14 - 0.03	0.19	278
CAN	0.66	0.61 - 0.71	0	-0.24	-0.32 - -0.16	0	-0.05	-0.11 - 0.02	0.15	-0.04	-0.10 - 0.03	0.3	341
HK (Non-White)	0.52	0.44 - 0.59	0.67	-0.29	-0.43 - -0.15	0	-0.05	-0.16 - 0.06	0.36	-0.11	-0.24 - 0.02	0.1	194
HK (White)	0.25	0.04 - 0.46	0.02	-0.4	-0.72 - -0.08	0.02	0.05	-0.49 - 0.59	0.85	-0.4	-0.87 - 0.07	0.09	20
NZ	0.6	0.53 - 0.66	0	-0.15	-0.25 - -0.05	0	-0.03	-0.11 - 0.05	0.47	-0.04	-0.12 - 0.05	0.38	243
SA	0.66	0.60 - 0.72	0	-0.26	-0.37 - -0.14	0	-0.06	-0.15 - 0.03	0.17	-0.02	-0.12 - 0.07	0.64	263
SG	0.5	0.44 - 0.57	0.95	-0.37	-0.49 - -0.25	0	-0.05	-0.15 - 0.04	0.28	-0.11	-0.22 - 0.00	0.06	239
UK	0.7	0.65 - 0.75	0	-0.27	-0.36 - -0.18	0	-0.04	-0.10 - 0.03	0.24	-0.04	-0.11 - 0.04	0.36	284
US	0.65	0.62 - 0.68	0	-0.18	-0.23 - -0.14	0	-0.05	-0.09 - -0.01	0.01	-0.02	-0.06 - 0.02	0.38	1,070
US, Weighted	0.65	0.62 - 0.68	0	-0.19	-0.23 - -0.14	0	-0.06	-0.09 - -0.02	0.01	-0.05	-0.09 - -0.00	0.03	1,070

Table S8 shows that the results are similar across a variety of vulnerable populations, those who answered all the scenario comprehension questions correctly, and various politically relevant subgroups. As described above, we also examine a racial subgroup of non-white in Australia, Canada, New Zealand, the United Kingdom, and the United States, and as both non-Chinese and non-white in Hong Kong and Singapore. The not college educated subgroup was not pre-specified, but is shown to demonstrate that our conclusions are not driven by the fact that our sample is slightly more educated than the general population.

Table S8: Study 2 - Main Results By Demographic Groups

	Prefers Scientists Conduct Integrated Study (0/1)			How Ethical (1-4): Integrated Minus Standard			How Scientifically Valid (1-4): Integrated Minus Standard			Would Take Vaccine (1-4): Integrated Minus Standard			N
	Mean	95% CI	p	Mean	95% CI	p	Mean	95% CI	p	Mean	95% CI	p	
All Participants	0.63	0.61 - 0.65	0	-0.24	-0.27 - -0.21	0	-0.05	-0.07 - -0.03	0	-0.05	-0.08 - -0.03	0	2,932
Vulnerable Populations													
Age 65+	0.74	0.70 - 0.78	0	-0.12	-0.18 - -0.06	0	-0.01	-0.06 - 0.04	0.67	0.04	-0.02 - 0.10	0.17	507
Essential Worker	0.57	0.54 - 0.60	0	-0.23	-0.28 - -0.17	0	-0.07	-0.12 - -0.02	0	-0.04	-0.10 - 0.01	0.1	880
Non-White (In HK & SG, Non-Chinese)	0.55	0.52 - 0.59	0	-0.29	-0.35 - -0.22	0	-0.09	-0.15 - -0.03	0	-0.1	-0.17 - -0.04	0	728
US: County cases >median	0.64	0.60 - 0.68	0	-0.19	-0.26 - -0.13	0	-0.06	-0.12 - -0.00	0.04	-0.04	-0.10 - 0.02	0.21	550
Political Groups (US Only)													
US: Republican	0.68	0.63 - 0.73	0	-0.14	-0.21 - -0.06	0	-0.03	-0.09 - 0.04	0.43	-0.05	-0.12 - 0.02	0.19	358
US: Democrat	0.63	0.58 - 0.68	0	-0.15	-0.23 - -0.07	0	-0.04	-0.11 - 0.03	0.26	-0.01	-0.07 - 0.06	0.85	331
US: Indep/Other	0.65	0.60 - 0.69	0	-0.27	-0.35 - -0.19	0	-0.1	-0.17 - -0.03	0.01	-0.08	-0.15 - -0.00	0.04	381
Robustness Checks													
Correctly answered all comprehension questions	0.68	0.66 - 0.70	0	-0.29	-0.33 - -0.26	0	-0.08	-0.10 - -0.05	0	-0.05	-0.08 - -0.02	0	1,745
No college degree	0.64	0.61 - 0.66	0	-0.2	-0.24 - -0.15	0	-0.03	-0.06 - 0.00	0.08	-0.04	-0.07 - 0.00	0.05	1,496

Table S5 shows the results are consistent across parameters we used in describing the integrated and standard trials. Unsurprisingly, respondents were slightly less likely to favor the integrated design when greater numbers of study participants receive the vaccine before a safety or immunogenicity problem with the vaccine could be detected.

Table S9: Study 2 - Main Results By Randomized Description of Trial Designs

	Prefers Scientists Conduct Integrated Study (0/1)			How Ethical (1-4): Integrated Minus Standard			How Scientifically Valid (1-4): Integrated Minus Standard			Would Take Vaccine (1-4): Integrated Minus Standard			N
	Mean	95% CI	p	Mean	95% CI	p	Mean	95% CI	p	Mean	95% CI	p	
All Participants	0.63	0.61 - 0.65	0	-0.24	-0.27 - -0.21	0	-0.05	-0.07 - -0.03	0	-0.05	-0.08 - -0.03	0	2,932
Speed of Standard Trial; How Many Months Integrated Accelerates Timeline													
12; Integrated 2 Mo Faster	0.57	0.53 - 0.61	0	-0.23	-0.30 - -0.16	0	-0.04	-0.09 - 0.01	0.15	-0.13	-0.19 - -0.06	0	501
12; Integrated 4 Mo Faster	0.63	0.59 - 0.68	0	-0.25	-0.32 - -0.18	0	-0.03	-0.08 - 0.03	0.36	-0.07	-0.13 - -0.00	0.04	495
12; Integrated 6 Mo Faster	0.69	0.65 - 0.73	0	-0.21	-0.29 - -0.14	0	-0.04	-0.10 - 0.02	0.16	-0.02	-0.08 - 0.04	0.55	509
18; Integrated 2 Mo Faster	0.58	0.53 - 0.62	0	-0.27	-0.34 - -0.19	0	-0.05	-0.11 - 0.02	0.14	-0.04	-0.11 - 0.04	0.33	451
18; Integrated 4 Mo Faster	0.64	0.60 - 0.68	0	-0.25	-0.32 - -0.18	0	-0.03	-0.09 - 0.03	0.33	-0.04	-0.10 - 0.02	0.19	490
18; Integrated 6 Mo Faster	0.65	0.60 - 0.69	0	-0.21	-0.28 - -0.14	0	-0.11	-0.17 - -0.05	0	-0.03	-0.09 - 0.03	0.33	486
# Of Integrated Subjects Who Get Vaccine Before Trial Stopped, if Vaccine Unsafe or Not Immunogenic													
250	0.67	0.64 - 0.70	0	-0.2	-0.25 - -0.15	0	-0.02	-0.06 - 0.02	0.35	-0.02	-0.07 - 0.02	0.31	992
400	0.61	0.58 - 0.64	0	-0.24	-0.29 - -0.19	0	-0.04	-0.08 - 0.01	0.09	-0.05	-0.09 - -0.00	0.05	943
1,000	0.6	0.57 - 0.63	0	-0.27	-0.32 - -0.21	0	-0.09	-0.13 - -0.05	0	-0.09	-0.14 - -0.04	0	997

Multivariate Regression Results

Figure S5 also visualizes coefficients from a multivariate regression predicting preference for the integrated trial. Variables denoted with (Std.) have been rescaled to standard deviation 1. Note that these coefficients should be interpreted as descriptive, not causal, and that all coefficients are calculated holding constant the other variables shown. The intercept of the regression is 0.64 and the omitted base categories in the regressions are: United States; age 18-24; male; no college degree; employment one of homemaker, student, unable to work, or unemployed since before the coronavirus pandemic; the integrated trial being study A; and average values of the rescaled variables. The coefficients below represent estimated differences from these base categories, holding constant the other variables in the model.

Figure S6 shows the results of a multivariate regression regressing the same variables as above on a binary indicator for whether respondents got all the comprehension questions about the integrated study correct. Gender, age, views towards vaccines, and scientific knowledge generally predict comprehension. This regression was not pre-registered.

Sensitivity Analysis

We also conduct a sensitivity analysis to determine how unrepresentative our sample would need to be in order to alter our conclusions. In particular, we compute [5]’s γ for the proportion that prefer the integrated study as 1.61 at a p-value of 0.10. This means that our statistical significance for the null hypothesis that integrated trials are equally preferred to standard trials would only no longer reach significance at the 0.10 level were individuals who do not prefer integrated trials to be 1.61 times less likely than individuals who prefer integrated trials to be selected for the survey. For comparison, it is very unusual in social science studies to observe a γ value greater than 2 [4].

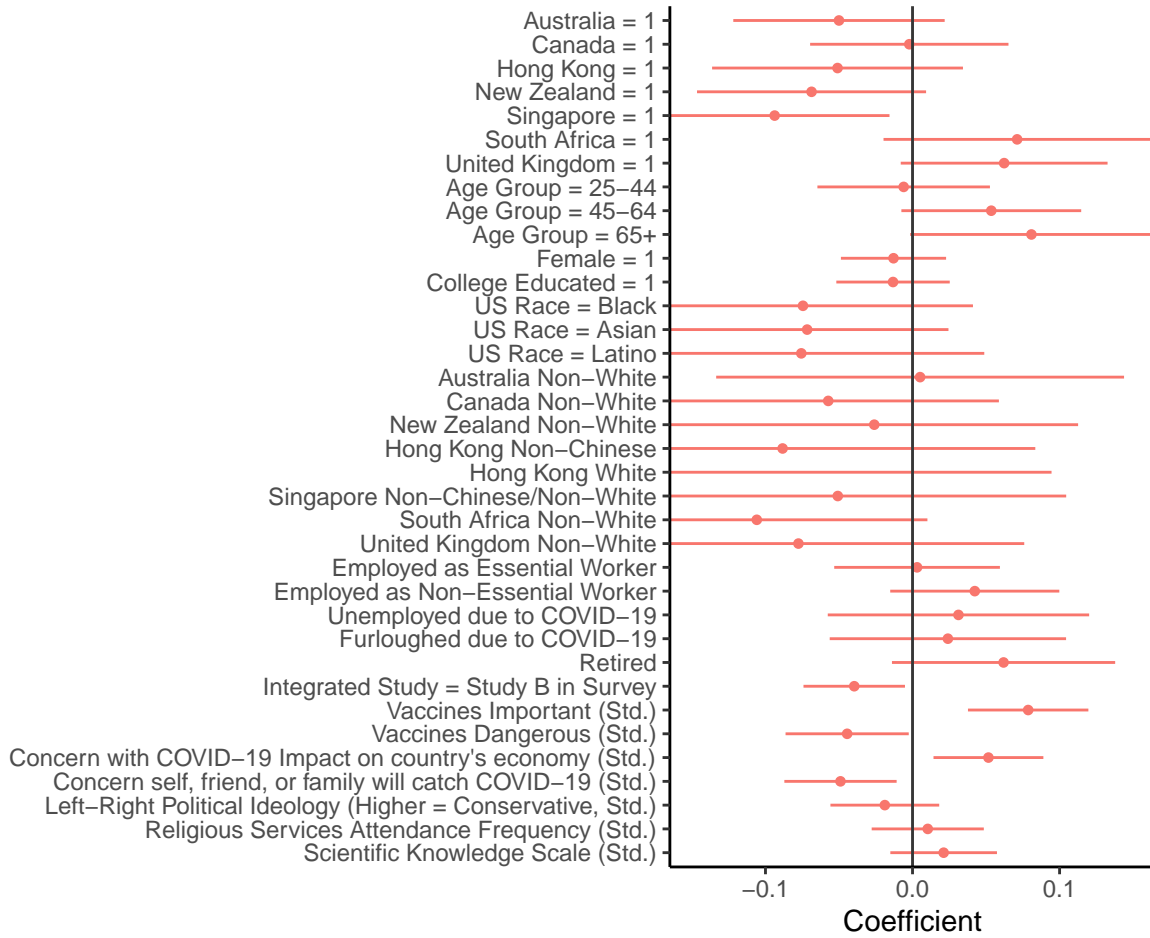


Figure S5: Predictors of Preference for Integrated Trial

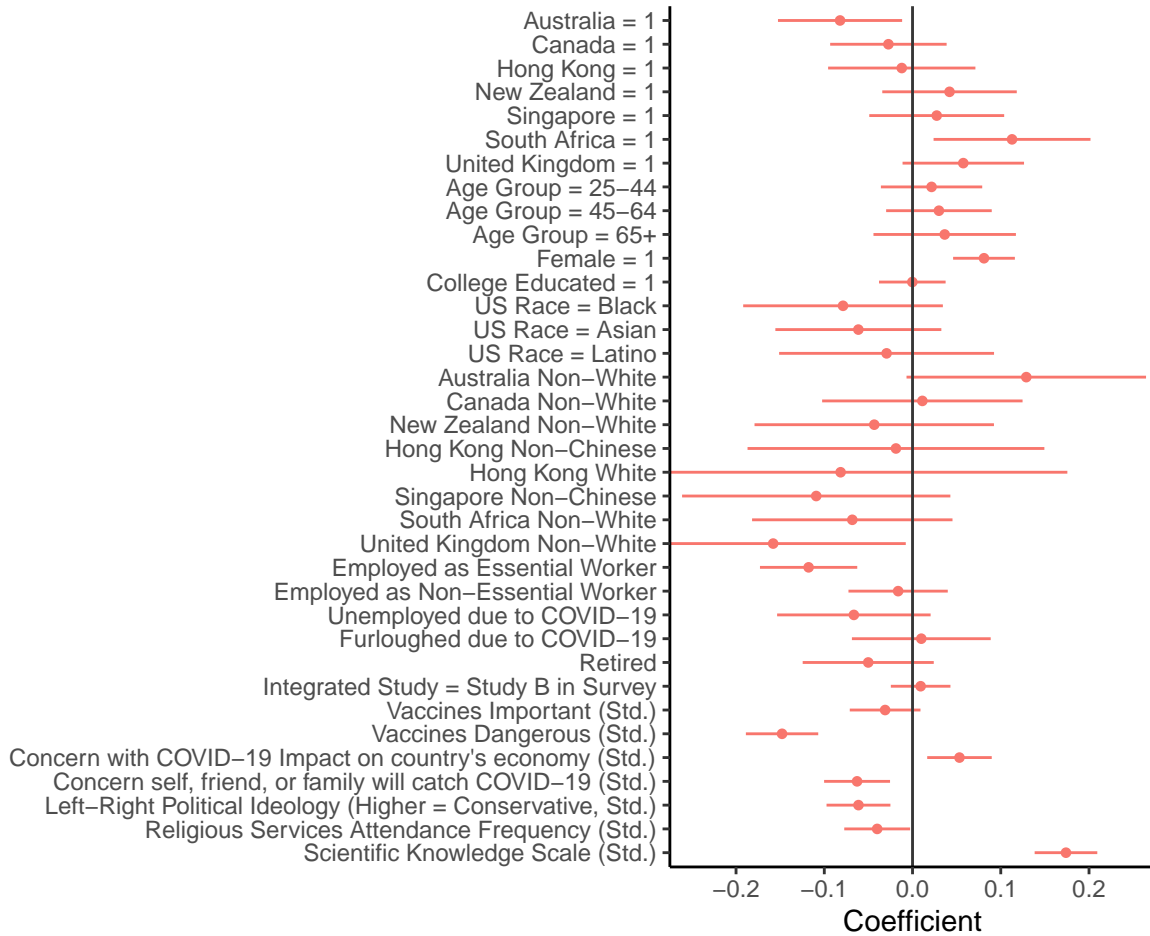


Figure S6: Predictors of Answering Scientific Knowledge Questions Correctly for Integrated Trial

Pre-Registration

Departures from Pre-Registration

We made only minor departures from our pre-registered analysis in our pre-analysis plan:

- Given that participants to the survey were more likely to be college educated than the general population, we present results for “non-college educated” participants as a subgroup to show our results are robust among this group, which we had not pre-specified.
- There were an unanticipatedly large number of participants in Hong Kong who identified as white, and these participants gave distinctive responses to all the questions, affecting our average characterization of Hong Kong. The 2016 Hong Kong Census estimates that only 0.8% of the Hong Kong population identifies as white [1], so we report the results of self-identified non-white and white Hong Kong participants separately. As Tables S2 and S6 show, self-identified whites in Hong Kong were especially unlikely to understand the scenarios correctly, suggesting this group of self-identified white Hong Kong residents may have been a subset of participants in Hong Kong who were answering the survey carelessly, including the racial identification question itself. Consistent with this interpretation, the median self-identified white participant in Hong Kong spent only 95 seconds reading the main study table (we did not allow participants to move on from the page describing the trial design until at least 60 seconds had gone by), versus a median of 124 seconds in the rest of the Hong Kong sample and a median of 181 in the sample outside of Hong Kong.

Pre-Registration Document

Our pre-registration document appears on the following pages. It was filed prior to the collection of our survey data.

**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 is 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.

References

- [1] Census and Statistics Department, Government of the Hong Kong Special Administrative Region of the People's Republic of China. Main tables, table a104. Available at <https://www.byccensus2016.gov.hk/en/bc-mt.html>.
- [2] Douglas R Green. Sars-cov2 vaccines: Slow is fast. *Science Advances*, 2020.
- [3] Jens Hainmueller. Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, 20(1):25–46, 2012.
- [4] Luke Keele. An overview of rbounds: An r package for rosenbaum bounds sensitivity analysis with matched data. *White Paper*, 2010.
- [5] Paul R Rosenbaum. Observational studies. 2020.
- [6] Robert Verity, Lucy C Okell, Ilaria Dorigatti, Peter Winskill, Charles Whittaker, Natsuko Imai, Gina Cuomo-Dannenburg, Hayley Thompson, Patrick GT Walker, Han Fu, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet Infectious Diseases*, 2020.
- [7] World Health Organization. Key criteria for the ethical acceptability of covid-19 human challenge studies. Technical report, World Health Organization, 2020.