

The effectiveness of moral messages on public health behavioral intentions during the COVID-19 pandemic

Jim A.C. Everett¹, Clara Colombatto², Vladimir Chituc², William J. Brady², & Molly J. Crockett²

¹Department of Psychology, University of Kent

²Department of Psychology, Yale University

Please address correspondence to J.A.C.Everett@kent.ac.uk or mj.crockett@yale.edu

Summary

With the COVID-19 pandemic threatening millions of lives, changing our behaviors to prevent the spread of the disease is a moral imperative. Here, we investigated the effectiveness of messages inspired by three major moral traditions on public health behavioral intentions. A sample of US participants representative for age, sex and race/ethnicity (N=1032) viewed messages from either a leader or citizen containing deontological, virtue-based, utilitarian, or non-moral justifications for adopting social distancing behaviors during the COVID-19 pandemic. We measured the messages' effects on participants' self-reported intentions to wash hands, avoid social gatherings, self-isolate, and share health messages, as well as their beliefs about others' intentions, impressions of the messenger's morality and trustworthiness, and beliefs about personal control and responsibility for preventing the spread of disease. Consistent with our pre-registered predictions, deontological messages had modest effects across several measures of behavioral intentions, second-order beliefs, and impressions of the messenger, while virtue-based messages had modest effects on personal responsibility for preventing the spread. These effects were observed for messages from leaders and citizens alike. Our findings are at odds with participants' own beliefs about moral persuasion: a majority of participants predicted the utilitarian message would be most effective. We caution that these effects are modest in size, likely due to ceiling effects on our measures of behavioral intentions and strong heterogeneity across all dependent measures along several demographic dimensions including age, self-identified gender, self-identified race, political conservatism, and religiosity. Although the utilitarian message was the least effective among those tested, individual differences in one key dimension of utilitarianism—impartial concern for the greater good—were strongly and positively associated with public health intentions and beliefs. Overall, our preliminary results suggest that public health messaging focused on duties and responsibilities toward family, friends and fellow citizens will be most effective in slowing the spread of COVID-19 in the US. Ongoing work is investigating whether deontological persuasion generalizes across different populations, what aspects of deontological messages drive their persuasive effects, and how such messages can be most effectively delivered across global populations.

Introduction

The COVID-19 pandemic threatens millions of lives. Compared to the seasonal flu, current estimates suggest that the novel coronavirus is nearly twice as contagious,¹ and it takes much longer for symptoms to present.² Thus, someone who feels otherwise healthy may be unknowingly transmitting the disease to others well before they even know they are sick. COVID-19 presents a growing public health crisis, and current worst-case estimates, by the Centers for Disease Control and Prevention and others, suggest that anywhere between 60% and 80% of people might eventually contract COVID-19, potentially killing tens of millions (Fink, 2020, *The New York Times*; Ferguson et al., 2020). Accordingly, any measures to prevent the transmission of the disease and avoid these worst-case scenarios are of great importance, and public health experts believe that “individual behaviour will be crucial to control the spread of COVID-19” (Anderson et al. 2020). These behaviors include washing hands more frequently and for longer, avoiding public gatherings, cancelling travel plans, and keeping distance from others. If measures like these are adopted, experts suggest that the deaths caused by coronavirus would drop dramatically (Ferguson et al. 2020).

Many people, from politicians to celebrities to health officials to local businesses, are urging members of the public to adopt these behaviors immediately as a moral imperative, and there are several distinct moral principles that may be persuasive. One might, for example, appeal to *utilitarian* moral theories (see e.g. Mill, 1863; Singer, 1972), focused on the expected consequences of failing to slow the spread of the pandemic. One might appeal to *deontological* moral theories (see e.g. Kant, 1785/1959; Scanlon, 2003), focused on notions of the duties, obligations, and responsibilities we have to our families and communities. Or one might appeal to moral theories based on notions of *virtue* and moral exemplars (see e.g. Aristotle, 350BC/2019; Hursthouse, 1999), considering what an ideal person would do during this crisis.

To our knowledge, there is almost no work that directly investigates how these different moral justifications might influence intentions to act, and certainly not in the context of the current pandemic. Past work in our lab and others suggests that deontological appeals will be more effective than utilitarian appeals, as deontological agents are seen as more trustworthy than utilitarian agents (Bostyn & Roets, 2017, Everett, Pizarro & Crockett 2016; Rom, Weiss & Conway 2017; Everett et al. 2018; Rom & Conway 2018) and people are more likely to be influenced by and accept advice from people they trust (Koenig, Clément, & Harris, 2004; Snizek & Van Swol, 2001; Van Swol, 2011). Other work on moral exemplars suggests that appeals based on virtue may also be effective in promoting individual prosocial behaviors (Han

¹ One study found that the median reproductive index (R_0) of the seasonal flu was 1.28 (Biggerstaff et al., 2014, *BMC Infectious Diseases*), while estimates from China suggest that R_0 for coronavirus is 2.2 (Li et al., 2020, *The New England Journal of Medicine*). Put another way, the average person with the flu will infect about one other person, while the average person with coronavirus will infect about two.

² According to the CDC (2018), the incubation period for the seasonal flu is about 2 days, while current estimates suggest that the incubation period for COVID-19 is up to two weeks (median is 5 days; Li et al., 2020, *The New England Journal of Medicine*).

et al. 2017), since upward moral comparison increases behavioral intentions to act prosocially (Zhang et al., 2017).

A second question concerns whether the effectiveness of moral messaging depends on the source of the message. Right now, people are getting information about coronavirus from their family, friends, strangers on social media, and people in leadership positions. Previous work suggests that the effectiveness of different moral justifications may depend on who communicates them. For example, people who follow utilitarian principles, spending their free time or extra money on needy strangers rather than their immediate family, are judged to be worse spouses but not worse political leaders (Everett et al. 2018). Even if, as we predict, deontological or virtue-based messages will be more effective than utilitarian messages, it might be that this effect is attenuated or even reversed when it comes from a person in authority, who is supposed to act more impartially and make cost/benefit decisions for the greater good.

Of course, not all people are reacting to this pandemic in the same way. In the US, there is evidence that demographic factors such as age and political identification are related to feelings and behaviors about COVID-19, with younger people and more conservative people feeling less concerned and reporting fewer changes to their behavior (YouGov, 2020). We therefore sought to investigate how demographic factors predict perceptions of pandemic threat and public health behaviors, and whether moral messaging effects are robust to controlling for these factors.

Finally, we examined how individual differences in utilitarianism relate to pandemic-related beliefs and intentions. Though much work in moral psychology explores utilitarianism in the context of sacrificial moral dilemmas, where someone must commit harm in service of the greater good (see e.g. Kahane et al., 2018; Everett & Kahane, 2020), utilitarian theories in ethics are most often concerned with an impartial concern for the wellbeing of all (see e.g. Singer, 2011). Thus, recent work has explored individual differences in these two components of utilitarianism separately—what Kahane, Everett and colleagues (2018) call *Instrumental Harm* (IH) and *Impartial Beneficence* (IB), respectively. Given the global nature of this pandemic, and the fact that public health behaviors can help not just one's own family and friends but also distant strangers, we anticipated that participants higher in IB, but not IH, would show greater intentions to adopt public health behaviors. Because past work has shown that mistrust in utilitarian (relative to deontological) agents is lower for people with utilitarian tendencies themselves, we expected that the effects of message type would be weaker for people with stronger utilitarian tendencies.

To test these hypotheses, we recruited a sample of US participants representative for the US population on age, sex and race/ethnicity (final $N = 1032$). Data were collected between 3/15/2020 and 3/16/2020. In a 2 x 4 between-subjects design, we randomly assigned participants to read a Facebook message from either a leader or citizen urging people to stay at home to prevent the spread of the coronavirus, providing either no moral justification (non-moral control condition) or a moral justification (deontological, virtue-based, or utilitarian; see **Table 1**). Following this, they answered questions about their own intentions to adopt public health

behaviors (washing hands, avoiding social gatherings, self-isolating, and sharing public health messages). Because we expected potential ceiling effects for behavioral intentions due to the strong social desirability of these behaviors, we also measured participants' second-order beliefs about the behavioral intentions of others, which we expected would be less sensitive to social desirability effects. Additionally, we measured perceptions of the morality and trustworthiness of the messenger and attitudes toward the pandemic, followed by measures of utilitarianism (Kahane, Everett et al. 2018) and demographics. Finally, we measured participants' own intuitions about whether deontological, virtue-based or utilitarian messages would be more effective in persuading others to take action. We pre-registered our hypotheses and methods (<https://osf.io/wy93v>) and posted all our materials and data (<https://osf.io/5vj8y/>) on OSF.

Table 1. Manipulation of moral messaging. Participants saw a Facebook post from either a citizen (high school teacher) or leader (Department of Education director) telling people to “STAY HOME, even if you don’t feel sick. Coronavirus is contagious even before you have symptoms.” This statement was followed by one of the messages in the table below.

Condition	Message
Non-moral control	We all need to do this, however difficult.
Deontological	We all need to do this, however difficult, because it is the right thing to do: it is our duty and responsibility to protect our families, friends, and fellow citizens. IT'S YOUR DUTY.
Virtue	We all need to do this, however difficult, because it's what a good person would do. Think about people you admire morally – what would they do? BE A GOOD PERSON.
Utilitarian	We all need to do this, however difficult, because these sacrifices are nothing compared to the much worse consequences for everyone if we carry on as usual. THINK OF THE CONSEQUENCES.

Results: Moral message effects

First, as specified in our pre-registration, we conducted ANOVAs to analyze the effects of moral message type (deontological vs. virtue-based vs. utilitarian vs. control), source (citizen vs. leader), and their interaction on behavioral intentions, second order beliefs, perceptions of the messenger, and pandemic attitudes.

Message source did not significantly affect most of our dependent measures. One important exception was intentions to share the message: participants reported stronger intentions to

share messages from leaders than citizens ($F(1,1023) = 5.21, p = .02$). Source did not interact with message type for any of our dependent measures, contrary to our predictions. In what follows, we focus on the effects of message type. Across several of our dependent measures, we find some modest effects of deontological and virtue-based messaging, in line with our predictions (**Fig. 1 & 2**).

We also observed strong relationships between demographic factors and most of our dependent measures. Because of this, we supplemented our pre-registered analyses with exploratory ANCOVAs testing the effects of message type, source, and their interaction while controlling for age, self-identified gender, self-identified race, education, income, employment status, political ideology, and religiosity. These are reported after our pre-registered analyses.

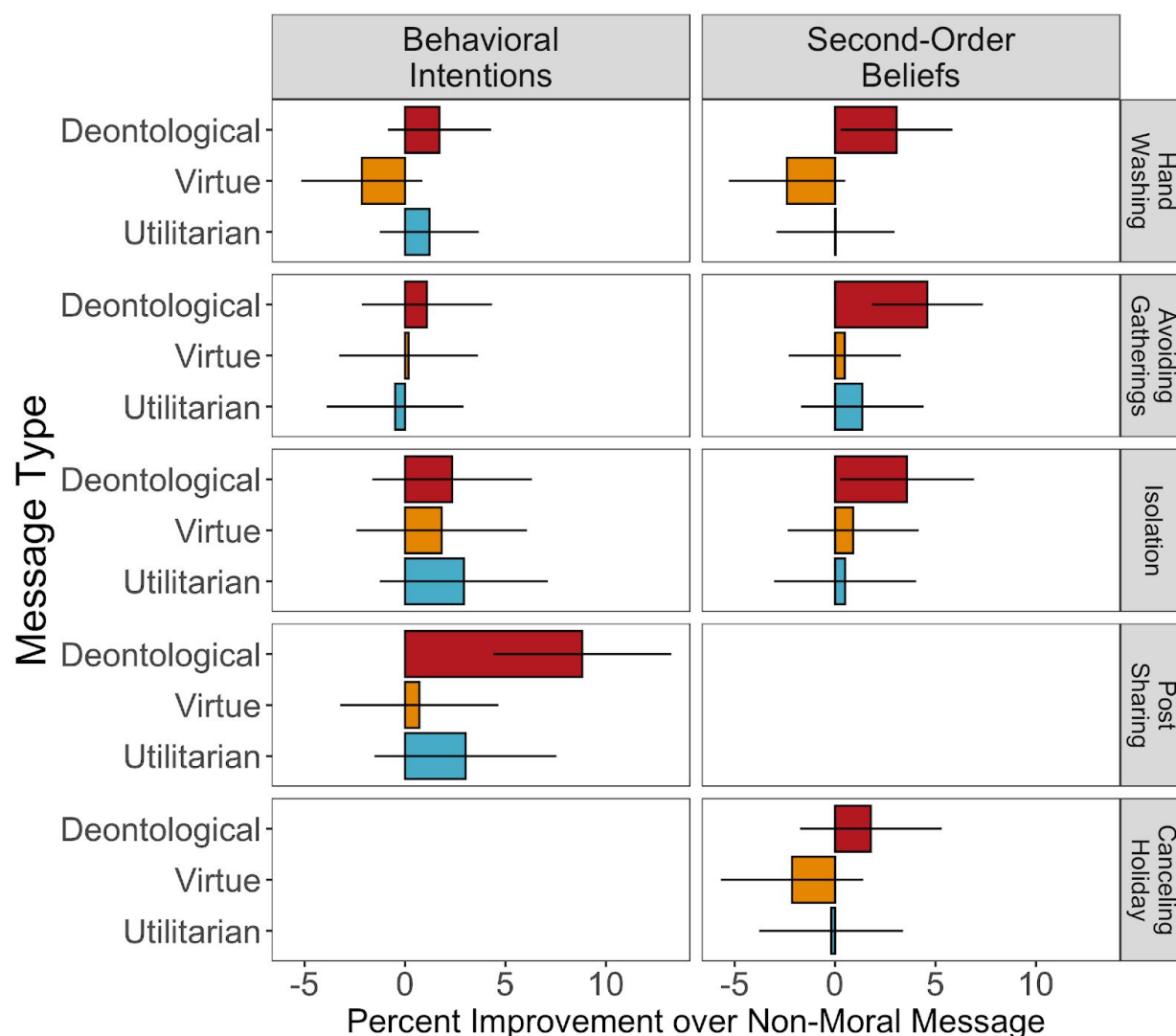


Figure 1. Effects of moral messaging on behavioral intentions and second-order beliefs about public health behaviors. Figure shows effects of deontological, virtue-based, and utilitarian messages relative to non-moral control as the percent change in behavioral intentions

(left panels) and second-order beliefs about others' intentions (right panels). Error bars represent 95% CIs.

Behavioral intentions. As expected, many of our self-reported behavioral intentions measures suffered from ceiling effects. Nevertheless, we found a main effect of message type on intentions to share the message ($F(3, 1023) = 3.95, p = .01$). Planned comparisons showed that participants reported significantly stronger intentions to share the deontological message ($M=3.34$) relative to the non-moral message ($M=2.82, t(1023)=3.10, p = .01$) and the virtue-based message ($M=2.86, t(1023)=2.83, p = .02$). All other pairwise comparisons were non-significant.

Next, we conducted exploratory ANCOVAs controlling for demographic factors. These additionally revealed a marginal effect of message type on intentions to wash hands ($F(3, 1007) = 2.29, p = .077$). Pairwise comparisons revealed a marginal difference such that deontological messages were more effective than virtue-based messages ($t(1005)=2.30, p = .1$). For intentions to share the message, there remained a significant effect of message type ($F(3, 1006) = 4.02, p = .007$), and pairwise comparisons again revealed significant differences between deontological messages and both the virtue-based ($t(1006)=2.82, p = .03$) and non-moral ($t(1006)=3.16, p = .01$) messages.

Second order beliefs. We observed a marginally significant main effect of message type on second-order beliefs about hand-washing ($F(3,1024) = 2.36, p = .07$). Planned comparisons revealed that participants reported significantly stronger second-order beliefs about handwashing after reading the deontological message ($M=4.95$) relative to the virtue-based message ($M=4.62, t(1024)=2.65, p = .04$). We also observed a marginal main effect of message type on second-order beliefs about avoiding social gatherings ($F(3,1024) = 2.24, p = .08$). Planned comparisons showed that the deontological message ($M = 4.68$) led to marginally stronger second-order beliefs about avoiding social gatherings, relative to the non-moral message ($M = 4.40, t(1024)=2.36, p = .09$).

In the ANCOVAs controlling for demographic factors, we also found a significant effect of message type on second-order beliefs about washing hands ($F(3, 1007) = 2.94, p = .03$), and pairwise comparisons revealed a significant difference such that deontological messages were more effective than virtue-based messages ($t(1007)=2.96, p = .02$). For second-order beliefs about avoiding social gatherings, we again found a marginally significant effect of message type ($F(3, 1007) = 2.25, p = .08$), and pairwise comparisons again revealed a marginal difference between deontological messages and non-moral messages ($t(1007)=2.37, p = .08$).

Perceptions of messenger. We observed a trend-level effect of message type on perceptions of messenger morality ($F(3,1024) = 2.40, p=0.07$). Planned comparisons showed that people who shared deontological messages were seen as significantly more moral ($M = 5.47$) than non-moral messengers ($M = 5.20; t(1024)=2.60, p = .05$). In the supplemental analyses at the end of this document, we report exploratory mediation analyses testing whether deontological

messages impact behavioral intentions and second-order beliefs via increasing perceived morality of the messenger.

Exploratory ANCOVAs controlling for demographic factors revealed no effects of message type on perceived morality or trustworthiness of the messenger, though in this analysis a significant effect of message source emerged for trustworthiness, such that participants reported higher trust in leaders compared to citizens ($F(3,1007) = 4.18, p = 0.04$).

Pandemic attitudes. We observed a trend-level effect of message type on feelings of personal responsibility for protecting others from contracting the disease ($F(3,1024) = 2.20, p=0.09$). Planned comparisons showed that virtue-based messages lead to a marginally greater sense of responsibility ($M = 5.63$) compared to non-moral messages ($M = 5.29$; $t(1024)=2.42, p = .07$).

In the ANCOVAs controlling for demographic factors, the effect of message type on feelings of personal responsibility was significant ($F(3,1007) = 2.70, p=0.045$), and planned comparisons revealed that virtue-based messages were significantly more effective than non-moral messages ($t(1007)=2.57, p = .05$).

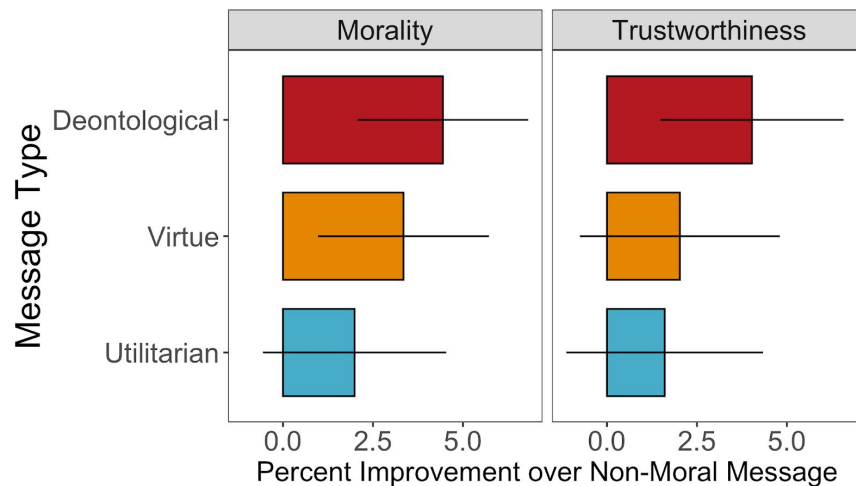


Figure 2. Impressions of messenger morality and trustworthiness from moral messages.

Figure shows effects of message conditions relative to non-moral control as the percent change in impressions of morality (left panel) and trustworthiness (right panel) of the messenger. Error bars represent 95% CIs.

Results: Lay beliefs about message effectiveness

To explore whether our findings reflect “common sense,” we explored our participants’ intuitions about whether deontological, virtue-based or utilitarian messages would be most effective.

Participants were presented with these three different moral messages as described in **Table 1**, and we asked them which would be most effective at reducing the spread of COVID-19. 53% of participants reported that they thought the utilitarian message would be most effective, 34%

thought the deontological message would be most effective, and 13% thought the virtue-based message would be most effective (one participant did not respond). We performed a chi-square test of independence and found that this difference was significant, $\chi^2(2, N = 1031) = 247.52, p < .001$, and post-hoc tests revealed that all pairwise comparisons were significant at $p < .001$. Thus, lay intuitions were at odds with what message was most effective.

Results: Demographic effects

In the tables below, we summarize main effects of demographic variables on our dependent measures from the exploratory ANCOVAs. Overall, older people and more religious people reported stronger intentions to practice public health behaviors and stronger beliefs about what others would do. People who self-identified as White, male and conservative reported weaker intentions to practice public health behaviors and weaker beliefs that others would do so. Education level was also negatively related to second-order beliefs.

Older people reported feeling more personally responsible for preventing the spread, and also felt people generally had more control over doing so. People who identified as Black and Asian also believed people had more control over social distancing. Meanwhile, people who identified as White and male reported feeling less personally responsible for preventing the spread. Political conservatives believed people had more control over practicing social distancing, while at the same felt less personally responsible for preventing the spread, and also held others less personally responsible.

Table 2. Demographic effects on behavioral intentions

	Washing hands more frequently	Avoiding gatherings	Self-isolation	Sharing message
Age	+++	+++	+++	++
Self-identified gender	--- (male)	-- (male)	-- (male)	
Self-identified race			--- (White)	--- (White)
Education	--		--	---
Income				
Employed status				
Political conservatism	---	---	---	---

Religiosity	+++			+++
-------------	-----	--	--	-----

“+” = positive relationship with variable, “-” = negative relationship with variable

“+++/--” = $p < 0.01$; “++/--” = $p < 0.05$; “+/-” = $p < 0.1$

Table 3. Demographic effects on second-order beliefs

	Washing hands more frequently	Avoiding gatherings	Self-isolation	Cancel vacation
Age	+++	+++	+++	+++
Self-identified gender				
Self-identified race	-- (White)	-- (White)	--- (White)	--- (White)
Education	-			
Income				
Employed status		++	+	++
Political conservatism			--	---
Religiosity	+	++	++	+

“+” = positive relationship with variable, “-” = negative relationship with variable

“+++/--” = $p < 0.01$; “++/--” = $p < 0.05$; “+/-” = $p < 0.1$

Table 4. Demographic effects on perceptions of messenger

	Morality	Trustworthiness
Age	+++	+++
Self-identified gender		
Self-identified race	++ (Asian)	
Education		
Income		
Employed status		
Political conservatism	---	---
Religiosity	+++	+++

“+” = positive relationship with variable, “-” = negative relationship with variable
“+++/-” = $p < 0.01$; “++/-” = $p < 0.05$; “+/-” = $p < 0.1$

Table 5. Demographic effects on pandemic attitudes

	Perceived control over social distancing	Own responsibility for preventing spread	Others' responsibility for preventing spread
Age	+++	+++	+
Self-identified gender		--- (Male)	- (Male)
Self-identified race	+++ (Black) + (Asian)	-- (White)	
Education			
Income			
Employed status			
Political conservatism	+++	---	---
Religiosity		+	

“+” = positive relationship with variable, “-” = negative relationship with variable
“+++/-” = $p < 0.01$; “++/-” = $p < 0.05$; “+/-” = $p < 0.1$

Results: Utilitarianism

Finally, we investigated the relationships between our dependent measures and utilitarianism scores, assessed using the Oxford Utilitarianism Scale (Kahane, Everett et al. 2018). The OUS dissociates two independent dimensions of utilitarian tendencies: impartial beneficence (IB; impartial concern for the greater good) and instrumental harm (IH; permissive attitude toward causing harm for the greater good). Across most of our dependent measures, we find consistent positive effects of IB, but not IH. We therefore focus here on the effects of IB. Below, we report the results from multiple regressions that include IB as a predictor along with demographic variables of age, self-identified gender, self-identified race, education, income, employment status, political ideology, and religiosity. We also ran analyses that allowed IB to interact with all three moral message types. Although IB strongly predicted most of our dependent measures, there were only a few interactions between IB and message type, none of which were predicted. Details of interactive effects can be found in the supplemental analyses at the end of this document.

Behavioral Intentions. Impartial beneficence had a positive effect on all of the self-reported behavioral intentions measures. Higher scores on impartial beneficence were associated with stronger intentions to wash hands more frequently ($\beta = .15$, $t(1007) = 4.10$, $p < .01$), stronger intentions to avoid social gatherings ($\beta = .26$, $t(1007) = 5.85$, $p < .01$), stronger intentions to completely self-isolate ($\beta = .39$, $t(1007) = 7.17$, $p < .01$), and stronger intentions to share the message ($\beta = .54$, $t(1007) = 9.84$, $p < .01$).

Second-order beliefs. Higher scores on impartial beneficence were associated with stronger beliefs that others would wash hands more frequently ($\beta = .14$, $t(1007) = 3.67$, $p < .01$), avoid social gatherings ($\beta = .21$, $t(1007) = 5.32$, $p < .01$), completely self-isolate ($\beta = .31$, $t(1007) = 7.06$, $p < .01$), and be willing to cancel a vacation they'd already paid for ($\beta = .13$, $t(1007) = 2.71$, $p < .01$).

Perceptions of messenger. Higher scores on impartial beneficence were associated with stronger perceptions of the messenger as more trustworthy ($\beta = .26$, $t(1007) = 7.28$, $p < .01$), and more moral ($\beta = .24$, $t(1007) = 7.40$, $p < .01$).

Pandemic attitudes. Higher scores on impartial beneficence were associated with stronger beliefs that people have control over practicing social distancing ($\beta = .18$, $t(1007) = 4.58$, $p < .01$), stronger feelings of personal responsibility for preventing the spread of the disease ($\beta = .35$, $t(1007) = 8.53$, $p < .01$), and stronger tendencies to view others as responsible for preventing the spread ($\beta = .33$, $t(1007) = 7.02$, $p < .01$).

Discussion

The best advice from public health experts emphasizes that, in order to slow the coronavirus pandemic, healthy people must take basic steps to change their behavior, and doing so has the potential to collectively save thousands if not millions of lives. This urgent moral message can fit into a number of ethical frameworks, and this study tested which might be most effective in a nationally representative U.S. sample.

We predicted that messages presented in a deontological framework, referencing the duties and obligations we have to one another, would increase behavioral intentions to wash hands, avoid gatherings, socially isolate, and share public health messages. We anticipated ceiling effects for these behavioral intentions, which are socially desirable (especially so now), and this seems to have been the case. We found a significant effect of deontological messages on intentions to share a public health message and a trend-level effect on intentions to wash hands. To reduce social desirability, and thus the likelihood of a ceiling effect, we also measured second-order beliefs about how likely *other* people reading a given message would be to change *their* behavior. Here, we found that participants reading deontological messages showed trends toward increasing second-order beliefs about hand washing and avoiding social gatherings. In line with our predictions, participants thought that people who posted deontological messages were more moral than those who posted other kinds of messages.

Based on past work, we also expected that virtue-based messages appealing to moral exemplars would be effective. For the most part, this was not the case, with the exception of feelings of personal responsibility. Here, virtue-based messages were significantly more effective than non-moral messages.

While theoretically based in previous work (e.g. Everett et al. 2016, 2018), our findings seem to be counterintuitive - at least to the participants in our study. When participants read the different moral framings of public health messages, the majority predicted that the utilitarian message would be most effective in convincing others to change their behavior. This was not the case, and it highlights the importance of empirically testing the effectiveness of different moral framings, especially since the stakes are now so high.

In line with past work, we expected that the source of the message, whether it came from a leader or a citizen, would interact with message type. We found no support for this prediction, and few effects of message source in general, though participants reported being more likely to share messages from leaders rather than citizens. It may be the case that our manipulation of leadership in this study (Department of Education director) was not perceived to have authority or expertise relevant to the pandemic crisis. Other types of leaders, such as public health officials, or high-power political figures may be more persuasive, and we will test this in future studies.

We observed strong heterogeneity across all dependent measures along several demographic dimensions including age, self-identified gender, self-identified race, political conservatism, and religiosity. In general, older people and more religious people reported stronger behavioral intentions and second-order beliefs about public health behaviors, while people who self-identified as White, male and conservative reported weaker behavioral intentions and second-order beliefs. Older people felt more personally responsible for preventing the spread of disease, while people who identified as White and male reported feeling less responsible. Political conservatives believed people had more control over preventing the spread while at the same time attributing lower responsibility to both themselves and others. In line with our predictions, the *Impartial Beneficence* component of Utilitarianism, but not *Instrumental Harm* (c.f. Kahane, Everett, et al. 2018), predicted behavioral intentions, beliefs about others' intentions, trust in the messenger, and perceptions of pandemic threat.

These results are, of course, limited in a number of important ways. We only tested one of many potential messages for each ethical framework. Kahane and Everett (2020) argue that we should adopt a multidimensional approach to utilitarianism, and it is plausible that messages related to different psychological instantiations of the utilitarian ideal may have different effects, just as different flavours of deontological justifications differentially influence person perception (Everett et al., 2016). Second, despite keeping the messages as structurally similar as possible, it remains possible that the messages unintentionally differed in ways other than their moral content. If this is the case, it is plausible that something apart from the deontological framing of

a message leads to increased effectiveness. That said, these concerns are allayed somewhat because that our participants actually expected the *utilitarian* message to be most effective at influencing others, suggesting that, on the face of it, participants did not perceive the deontological message as more persuasive.

On a related note, in conducting an online survey, we are inherently limited in exploring behavioral intentions and second-order beliefs, rather than actual behaviors to promote public health. These results are an important first step in exploring what types of messaging will be most effective at encouraging protective public health behaviors, but future work remains to determine whether these beliefs and intentions translate into actual changes in behavior.

Perhaps most importantly, our effects are small and in many cases do not pass conventional levels of statistical significance. However, in the context of a public health crisis, even changing the behavior of a small percentage of individuals can save lives, and small changes in framing of messages are a cheaply-implemented tool that could be readily implemented on communications platforms. We intend to replicate these findings in a larger sample as soon as possible. Future research also needs to test whether different types of messages are more effective for different segments of the population; whether our main results replicate across different countries; and how to most effectively deliver such messages, both in terms of the context of the message (e.g., embedded within a personal narrative, news story or generic public service announcement) and its source (e.g., whether it comes from someone familiar, a stranger, or someone with authoritative scientific or moral expertise). For the time being, however, as we search for evidence-based recommendations about how to persuade others to change their behavior to prevent the spread of COVID-19, our preliminary conclusion is that messages focused on duties and responsibilities toward family, friends and fellow citizens may be an effective strategy.

Methods

Participants. We aimed to recruit a sample of 1100 participants from Prolific Academic, representative of the US population on age, sex and race. Power analyses showed that to detect a small effect size ($f = 0.10$) with 90% power, we needed a final sample of 1053 participants. We expected our manipulation to produce small effect sizes on our dependent measures. Nevertheless, these small effects are incredibly important to know in this pandemic context. With the exponential growth of COVID-19, even a small fraction of the population changing their behavior to prevent the spread can save lives, just as political targeting affecting less than 1% of the sample can have drastic effects on voting and, ultimately, electoral outcomes (Bond et al. 2012, *Nature*). We ultimately recruited 1102 participants. As specified in our pre-registration, we excluded 58 participants for failing comprehension checks, and while not specified in our pre-registration, we also excluded a further 12 participants who completed the survey more than once. This left a final sample of 1032 participants.

Experimental design. In a 2x2 between-subjects design, participants were randomly assigned to read a Facebook message urging people to stay home even if they don't have symptoms. Between subjects, we varied the source of the message (leader vs citizen) and the content of that message (deontological vs virtue ethics vs utilitarian vs non-moral control). See <https://osf.io/5vj8y/> for full survey materials, including screenshots of stimuli. After providing informed consent, participants read the Facebook message and answered two multiple-choice comprehension check questions requiring them to correctly recall the job role of the poster and the moral justification they used to encourage others to stay home. Participants then answered a series of questions about their intentions and beliefs relating to the ongoing pandemic crisis, followed by demographic and individual difference measures, as detailed below.

Dependent measures.

Behavioral Intentions

In the first set of DVs, participants indicated their intentions to engage in protective-behavior relating to COVID-19. We included three such questions, from comparatively low effort to comparatively high effort.

- After reading this message, how likely is it that you will always wash your hands whenever you enter work or come home, for at least 2 weeks, even if you don't feel sick? (1 = *extremely unlikely*, 7 = *extremely likely*)
- After reading this message, how likely is it that you will avoid public gatherings for at least 2 weeks, even if you don't feel sick? (1 = *extremely unlikely*, 7 = *extremely likely*)
- After reading this message, how likely is it that you will stay at home and avoid all social contact for at least 2 weeks, even if you don't feel sick? (1 = *extremely unlikely*, 7 = *extremely likely*)

Next, we asked participants to indicate “How likely is it that you would share this post to your own social media?” (1 = *extremely unlikely*, 7 = *extremely likely*)

Second-Order Beliefs

In the second set of DVs, we had participants indicate how much they thought other people would engage in the same protective behaviors if they had also read this message, with an additional item about cancelling a vacation.

- How likely is it that someone else who reads this would, for the next 2 weeks, always wash their hands whenever they enter work or come home, even if they don't feel sick? (1 = *extremely unlikely*, 7 = *extremely likely*)
- How likely is it that someone else who reads this would avoid public gatherings for at least 2 weeks, even if they don't feel sick? (1 = *extremely unlikely*, 7 = *extremely likely*)

- How likely is it that someone else who reads this will stay at home and avoid all social contact for at least 2 weeks, even if they don't feel sick? (1 = *extremely unlikely*, 7 = *extremely likely*)
- How likely is it that someone else who reads this would cancel an upcoming international vacation they've already paid for, even if they don't feel sick? (1 = *extremely unlikely*, 7 = *extremely likely*)

In this same category, though it was asked at the end of the survey, we asked participants to look at all three moral messages and rate which they thought would be most convincing:

- “If you were trying to convince someone to take action to reduce the spread of COVID-19 (coronavirus), which of these messages do you think would be most effective? (*forced choice of the wording from the three moral manipulations: see Table 1*)

Perceptions of Messenger

Next, participants completed two measures evaluating the messenger.

- Think about the person who posted this message, Nicholas Miller. How moral, or good a person do you think they are? (1 = *very immoral/bad*, 7 = *extremely moral/good*)
- Think about the person who posted this message, Nicholas Miller. How trustworthy do you think they are? (1 = *not trustworthy at all*, 7 = *extremely trustworthy*)

Pandemic Attitudes

Participants were then asked to answer some questions about the COVID-19 pandemic specifically. Some of these were intended predictor variables, and some as outcome variables. Predictor, or individual difference, measures are described below. As an outcome, or dependent measure, we asked participants:

- How responsible do you feel about your own role in preventing further outbreak of COVID-19? (1 = *not at all responsible*, 7 = *very responsible*)
- If someone does NOT stay at home and avoid social contact for the next 2 weeks, how responsible do you think they are for further outbreak of COVID-19? (1 = *not at all responsible*, 7 = *very responsible*)
- How much control do you think that people have over whether they stay at home and avoid social contact? (1 = *no control at all*, 7 = *complete control*)

Lay intuitions about message effectiveness

Participants were asked the following question:

“If you were trying to convince someone to take action to reduce the spread of COVID-19 (coronavirus), which of these messages do you think would be most effective?”

They selected between the following three options (utilitarian, deontological, and virtue-based):

1. We all need to do this, however difficult, because these sacrifices are nothing compared to the much worse consequences for everyone if we carry on as usual. THINK OF THE CONSEQUENCES.
2. We all need to do this, however difficult, because it is the right thing to do: it is our duty and responsibility to protect our families, friends, and fellow citizens. IT'S YOUR DUTY.
3. We all need to do this, however difficult, because it's what a good person would do. Think about people you admire morally – what would they do? BE A GOOD PERSON.

Demographics and individual differences.

We asked participants some questions about their beliefs on COVID-19 to use as predictor, or individual difference, measures. These were:

- How much of a threat do you think COVID-19 (coronavirus) is? (1 = not a threat at all; 7 = extremely threatening)
- How likely do you think it is that you yourself will die as a result of the COVID-19 pandemic? (1 = *extremely unlikely*, 7 = *extremely likely*)
- How likely do you think it is that someone you know will die as a result of the COVID-19 pandemic? (1 = *extremely unlikely*, 7 = *extremely likely*)
- How unpleasant would it be for you personally to stay at home and avoid social contact for the next 2 weeks? (1 = *not at all unpleasant*; 7 = *extremely pleasant*)
- Over the past week, how often have you sought out information about the COVID-19 pandemic (e.g. news articles, video, social media posts, etc.)? (1 = *rarely*, 4 = *occasionally*, 7 = *frequently*)
- Are there currently any cases of COVID-19 in the city where you live? (yes, no, not sure)
- Has anyone you know personally tested positive for coronavirus?(yes, no, not sure)

All participants completed the Oxford Utilitarianism Scale (Kahane, Everett, et al. 2018). The scale consists of nine items in two subscales: instrumental harm (OUS-IH) and impartial beneficence (OUS-IB). The OUS-IB subscale consists of five items that measure endorsement of impartial maximization of the greater good, even at great personal cost (e.g., 'It is morally wrong to keep money that one doesn't really need if one can donate it to causes that provide effective help to those who will benefit a great deal'). The OUS-IH subscale consists of four items relating to willingness to cause harm so as to bring about the greater good (e.g., 'It is morally right to harm an innocent person if harming them is a necessary means to helping several other innocent people').

Finally, participants completed a series of demographic questions (see survey on OSF for full wording and description). Key demographic variables are depicted in Table 2.

Table 2. Demographic characteristics of the overall sample.

Demographic		Frequency	Percent
Self-identified gender (Female)		519	50.29%
Age		44.99 (15.97)	
Self-identified race	Asian	71	6.89%
	Black	133	12.90%
	Unspecified	79	7.66%
	White	748	72.55%
Education	High School	105	10.17%
	Some College	268	25.97%
	2-Year Degree	116	11.24%
	4 Year Degree	351	34.01%
	PostGrad/Professional	192	18.60%
Employment status	Full-Time	443	43.05%
	Unemployed	295	28.67%
	Student	84	8.16%
	Part-Time	207	20.12%
Political Affiliation	Republican	178	17.25%
	Democrat	512	49.61%
	Independent	256	24.81%
	None	56	5.43%
	Other	30	2.91%

Acknowledgments

We thank Sophie Harrington, Karissa McCright, and Lydia Schooler for assistance with survey building, Nicholas Christakis, Jon Jachimowicz and Crockett Lab members for fast and helpful

feedback on study design, and Prolific Academic for financial and technical support for data collection.

References

- Anderson, R. M., Heesterbeek, H., Klinkenberg, D., & Hollingsworth, T. D. (2020). How will country-based mitigation measures influence the course of the COVID-19 epidemic? *The Lancet*, 0(0). [https://doi.org/10.1016/S0140-6736\(20\)30567-5](https://doi.org/10.1016/S0140-6736(20)30567-5)
- Aristotle. (2019). *Nicomachean Ethics* (T. Irwin, Trans.). Hackett Publishing.
- Biggerstaff, M., Cauchemez, S., Reed, C., Gambhir, M., & Finelli, L. (2014). Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: A systematic review of the literature. *BMC Infectious Diseases*, 14(1), 480. <https://doi.org/10.1186/1471-2334-14-480>
- Bostyn, D. H., & Roets, A. (2017). Trust, trolleys and social dilemmas: A replication study. *Journal of Experimental Psychology: General*, 146(5), e1.
- CDC. (2018, August 27). How Flu Spreads. Centers for Disease Control and Prevention. <https://www.cdc.gov/flu/about/disease/spread.htm>
- Everett, J. A., Faber, N. S., Savulescu, J., & Crockett, M. J. (2018). The costs of being consequentialist: Social inference from instrumental harm and impartial beneficence. *Journal of Experimental Social Psychology*, 79, 200-216.
- Everett, J. A., & Kahane, G. (2020). Switching Tracks? Towards a Multidimensional Model of Utilitarian Psychology. *Trends in Cognitive Sciences*.
- Everett, J. A., Pizarro, D. A., & Crockett, M. J. (2016). Inference of trustworthiness from intuitive moral judgments. *Journal of Experimental Psychology: General*, 145(6), 772.
- Ferguson, N. M., Laydon, D., Nedjati-Gilani, G., Imai, N., Ainslie, K., Baguelin, M., Bhatia, S., Boonyasiri, A., Cucunubá, Z., Cuomo-Dannenburg, G., Dighe, A., Fu, H., Gaythorpe, K., Thompson, H., Verity, R., Volz, E., Wang, H., Wang, Y., Walker, P. G., ... Ghani, A. C. (2020). Impact of non-pharmaceutical interventions (NPIs) to reduce COVID- 19 mortality and healthcare demand. 20.
- Fink, S. (2020, March 13). Worst-Case Estimates for U.S. Coronavirus Deaths. *The New York Times*. <https://www.nytimes.com/2020/03/13/us/coronavirus-deaths-estimate.html>
- Han, H., Kim, J., Jeong, C., & Cohen, G. L. (2017). Attainable and Relevant Moral Exemplars Are More Effective than Extraordinary Exemplars in Promoting Voluntary Service Engagement. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.00283>

Hursthouse, R. (1999). *On virtue ethics*. OUP Oxford.

Kahane, G., Everett, J. A. C., Earp, B. D., Caviola, L., Faber, N. S., Crockett, M. J., & Savulescu, J. (2018). Beyond sacrificial harm: A two-dimensional model of utilitarian psychology. *Psychological Review*, 125(2), 131–164. <https://doi.org/10.1037/rev0000093>

Kant, I. (1959). *Foundation of the metaphysics of morals* (L. W. Beck, Trans.). Indianapolis, IN: Bobbs-Merrill. (Original work published 1785)

Koenig, M. A., Clément, F., & Harris, P. L. (2004). Trust in testimony: Children's use of true and false statements. *Psychological Science*, 15(10), 694-698.

Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., Ren, R., Leung, K. S. M., Lau, E. H. Y., Wong, J. Y., Xing, X., Xiang, N., Wu, Y., Li, C., Chen, Q., Li, D., Liu, T., Zhao, J., Li, M., ... Feng, Z. (2020). Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *The New England Journal of Medicine*. <https://doi.org/10.1056/NEJMoa2001316>

Mill, J. S. (1863). *Utilitarianism*. London, England: Parker, Son, and Bourne.

Rom, S. C., Weiss, A., & Conway, P. (2017). Judging those who judge: Perceivers infer the roles of affect and cognition underpinning others' moral dilemma responses. *Journal of Experimental Social Psychology*, 69, 44-58.

Rom, S. C., & Conway, P. (2018). The strategic moral self: Self-presentation shapes moral dilemma judgments. *Journal of Experimental Social Psychology*, 74, 24-37.

Scanlon, T.M., 2003, *The Difficulty of Tolerance: Essays in Political Philosophy*, Cambridge: Cambridge University Press.

Singer, P. (1972). Famine, affluence, and morality. *Philosophy & public affairs*, 229-243.

Singer, P. (2011). *The expanding circle: Ethics, evolution, and moral progress*. Princeton, NJ: Princeton University Press.

Snizek, J. A., & Van Swol, L. M. (2001). Trust and expertise in a judge advisor system. *Organizational Behavior and Human Decision Processes*, 84, 288–307.

Van Swol, L. M. (2011). Forecasting another's enjoyment versus giving the right answer: Trust, shared values, task effects, and confidence in improving the acceptance of advice. *International Journal of Forecasting*, 27(1), 103-120.

COVID-19 report: How the virus is affecting everything, from politics to brands | YouGov. (2020). Retrieved March 18, 2020, from <https://today.yougov.com/topics/science/articles-reports/2020/03/12/covid-19-report-how-virus-affecting-world-politics>

Zhang, H., Chen, S., Wang, R., Jiang, J., Xu, Y., & Zhao, H. (2017). How Upward Moral Comparison Influences Prosocial Behavioral Intention: Examining the Mediating Role of Guilt and the Moderating Role of Moral Identity. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.01554>

Supplemental Analyses

Mediation of messaging effects by impressions of messenger

We conducted a series of exploratory mediation analyses to investigate whether the significant effects of message type on our dependent measures were mediated by impressions of the messenger, in the form of increased trust and morality. In other words, might different types of moral justifications influence behavior and beliefs by specifically altering perceptions of the messenger's character? Because impressions of trustworthiness were not affected by message type, we focus here on impressions of morality, which as predicted were affected by the different types of justifications (see Results in Main Text). Across all the significant effects message type had on intentions, beliefs, and attitudes, we found that increased perceptions of messenger morality significantly mediated the effects of deontological and virtue-based messaging when the relevant contrasts involved non-moral messaging, but not when the relevant contrasts involved other types of moral messaging. We describe these results below. All confidence intervals were calculated via bootstrapping with 5,000 samples.

Behavioral intentions. Impressions of the messenger's morality mediated the effects of the deontological message on intentions to share the post relative to the non-moral message (indirect effect = .16, 95% CI = [.04; .30], $p = .01$), but not these same effects of deontological message relative to the virtue-based message (indirect effect = .04, 95% CI = [-.08; .17], $p = .51$).

Second order beliefs. Impressions of the messenger's morality did not mediate the effects of the deontological message compared to the virtue-based message on second-order beliefs about hand-washing (indirect effect = .02, 95% CI = [-.05; .09], $p = .53$). However, they did mediate the effects of the deontological message compared to the non-moral message on second-order beliefs about avoiding social gatherings (indirect effect = .09, 95% CI = [.02; .16], $p = .01$).

Pandemic attitudes. Impressions of the messenger's morality marginally mediated the effects of the deontological message on responsibility for protecting others from contracting the disease relative to the virtue-based message (indirect effect = .09, 95% CI = [.00; .18], $p = .05$).

Moderation of message effectiveness by impartial beneficence.

To explore whether impartial beneficence moderates the effects of our manipulation, we conducted a series of ANCOVAs that include IB as a predictor along with demographic variables of age, self-identified gender, self-identified race, education, income, employment status, political ideology, and religiosity and allowed IB to interact with message type. The strong effect of IB on our dependent measures is reported in the main text; we report here the interactions. Impartial beneficence did not interact with moral message type on behavioral intentions or perceptions of the messenger, but it had a trending influence on some aspects of

pandemic perception and second-order beliefs (full results are available at the OSF: <https://osf.io/5vj8y/>). First, the positive effects of deontological messaging and the negative effects of virtue-based messaging were stronger for individuals who scored higher on impartial beneficence. In particular, the positive effect of deontological messages on frequency of seeking information about the pandemic was weaker in people lower in impartial beneficence ($\beta = .25$, $t(1007) = 2.02$, $p = .04$), while the negative effect of virtue-based messages on beliefs that others would avoid gatherings was stronger in people higher in impartial beneficence ($\beta = -.20$, $t(1007) = -1.93$, $p = .05$), and so was the negative effect of virtue-based messages on beliefs that others would self-isolate ($\beta = -.30$, $t(1007) = -2.47$, $p = .01$).