I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13560

Socially Optimal Mistakes? Debiasing COVID-19 Mortality Risk Perceptions and Prosocial Behavior

Martin Abel
Tanya Byker
Jeffrey Carpenter

JULY 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13560

Socially Optimal Mistakes? Debiasing COVID-19 Mortality Risk Perceptions and Prosocial Behavior

Martin Abel
Middlebury College, JPAL and IZA

Tanya Byker
Middlebury College

Jeffrey Carpenter
Middlebury College and IZA

JULY 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Socially Optimal Mistakes? Debiasing COVID-19 Mortality Risk Perceptions and Prosocial Behavior*

The perception of risk affects how people behave during crises. We conduct a series of experiments to explore how people form COVID-19 mortality risk beliefs and the implications for prosocial behavior. We first document that people overestimate their own risk and that of young people, while underestimating the risk old people face. We show that the availability heuristic contributes to these biased beliefs. Using information about the actual risk to debias people's own risk perception does not affect donations to the Centers for Disease Control but does decrease the amount of time invested in learning how to protect older people. This constitutes a debiasing social dilemma. Additionally providing information on the risk for the elderly, however, counteracts these negative effects. Importantly, debiasing seems to operate through the subjective categorization of and emotional response to new information.

JEL Classification: C91, D91, H41

Keywords: risk perception, prosocial behavior, debiasing, experiment

Corresponding author:
Martin Abel
Middlebury College
14 Old Chapel Rd
Middlebury, VT 05753
USA
E-mail: mabel@middlebury.edu

* We thank Noah Whiting for outstanding research assistance. In addition, this paper greatly benefited from comments by participants of seminars at Rotterdam University and Middlebury College. The experiment was registered under registry number AEARCTR-0005579 and IRB approval was obtained from Middlebury College. All errors and omissions are our own. Declarations of interest: none.

1 Introduction

With news of the COVID-19 pandemic dominating every media outlet during the spring of 2020, there is reason to believe that people may have overestimated the mortality risk of the virus. As people are exposed to sickness and death at every click of the mouse, every button pressed on the television remote and every turn of the car radio dial, visceral instances of the disease will naturally accumulate and loom large at the top on one's mind. Not only does this news induce anxiety, it naturally makes serious instances of the disease more available for recall. While the stress induced by having the virus dominate one's thoughts is undoubtedly bad, it is less clear that overestimating the case mortality rate (CMR) of COVID-19, itself, is detrimental for society.

From a public policy perspective, heightened risk perceptions may induce people to adopt behavior to reduce the chances they will contract the virus such as following social distancing guidelines, washing their hands and wearing a mask. Importantly, many of these activities provide both private and public benefits. As individuals, based on their misperceptions of the risk, become more vigilant about hygiene, they coterminously create positive externalities that help reduce the risk to others. Misinformation may thus lead to socially optimal behavior, especially among younger people who face the lowest mortality risk but tend to overestimate the risk the most. This dilemma raises important policy questions. Does it make sense, from a public health perspective, to correct or *debias* risk perceptions during a pandemic? That is, does debiasing lead those who overestimate the risk to pull back on their efforts to protect themselves, inadvertently reducing the positive externalities they generate?[1]

We report evidence from a series of online experiments conducted in the U.S. during March and May, 2020 designed to assess how people, most of whom are middle-aged or younger, perceive the risk of COVID-19 and whether providing actual risk information alters their behavior. Specifically, we first elicited perceptions of the full COVID-19 age-mortality gradient and find that people do, indeed, systematically overestimate the mortality risk for young people, while simultaneously underestimating the risk for the elderly. In addition, we find that people living in areas with more COVID-19 victims and those who consume left-leaning media are more likely to overestimate the risks, suggesting that exposure to the crisis increases risk perceptions (Simonov et al., 2020; Pennycook et al., 2020). Lastly, we find that those who are more likely to be "intuitive thinkers," as indicated by the results of the cognitive reflection test (CRT), overestimate the risks to an even greater extent (similar to Pennycook et al. (2020)). Importantly, the CRT has recently been shown to be a good indicator of one's reliance on the use of decision-making heuristics (Frederick, 2005; Toplak et al., 2011; Johnson et al., 2014; Morsanyi et al., 2014).

[1] A large literature indicates that individual behavior, including prosocial behavior, is influenced by risk perceptions (e.g., Brewer et al. (2007); Gong (2015); Delavande and Kohler (2016)).

Potential bias and the use of heuristics to assess risk are the impetuses for our first experiment. The *availability heuristic* suggests that people tabulate risks based on instances that come to their minds easily (i.e., the instances that are more "available") and extrapolate linearly to the circumstances and risks of others, even when the actual risk gradients may be nonlinear (Tversky and Kahneman, 1973). If we can immediately recall multiple examples of something, like COVID-19 deaths, we believe they are very common. Previous work shows that instances become more available the more they appear in the media (Combs and Slovic, 1979; Wahlberg and Sjoberg, 2000; Agha, 2003; Romer et al., 2003; Kpanake et al., 2008; Frh, 2017; Slovic, 2000), the more vivid they are (Shedler and Manis, 1986; Bensi et al., 2003; Sjöberg and Engelberg, 2010; Dillard and Main, 2013), the more personal they are (Keller et al., 2006) and the more they induce emotional responses (Pachur et al., 2012; Sobkow et al., 2016).

Similar to Lichtenstein et al. (1978) who studied health risks more broadly, we find evidence that the availability heuristic influences COVID-19 risk perceptions. We see that experimentally imposing a cognitive load, a standard method to induce heuristic use, leads to increased risk perceptions, as predicted by availability and not the other common heuristics linked to risk perceptions. However, while also experimentally increasing the availability of victims did not have a significant effect on risk perceptions, the treatment effect was in the correct direction and likely muted due to COVID-19 fatalities already being salient to most participants. Supporting this broader correlation across our treatments, participants who knew of a person that died from COVID-19 are more likely to overestimate the risk, especially if the victim was part of their social network. These findings are consistent with much of the existing literature, documenting that heuristics - including availability - are more likely to be used when someone is stressed (Butler and Mathews, 1987; Shaham et al., 1992), as during a pandemic, or under cognitive load, which reduces the influence of analytical thought on choice (Schaeffer, 1989; Kassam et al., 2009; Heereman and Walla, 2011).

Given most people overestimate the risk to themselves of the virus, the natural question to ask is whether they can be debiased and, if so, what the effects are of correcting their risk perceptions on prosocial behavior. We conduct a second experiment assessing different ways to debias individual risk assessments. In one treatment, participants were informed of their own mortality risk (based on the best information at the time). In a second treatment, participants were informed of their own actual risk *and* the mortality risk of older individuals. Participants in a third treatment were provided the same information as those in the second treatment, but the information was tailored to a specific elderly people that respondents felt close to. The rationale for the third treatment was to test if an emotional connection could increase the chances that the information would be heeded (Traczyk et al., 2015).

Several recent studies on social distancing and other prosocial behavior during the pandemic raise concerns about the reliability of self-reported behavior data.[2] For this reason,

[2]For example, Falco and Zaccagni (2020) find that text message reminders about COVID-19 affect intended behavior but that intentions do not translate into actual changes in behavior.

we focused on observed and incentivized behavior in this second experiment. We observed, unbeknown to participants, how much time they invest watching informational videos related to the spread of COVID-19 and solicited donations from our participants to the Center for Disease Control's (CDC) COVID-19 emergency fund.

We find mixed results of debiasing. Learning just about your own mortality risk does not affect donations. That said, once we added the information about the actual risk to older people, which participants tend to underestimate, they are almost 7 percentage points (20%) more likely to make a donation. We further find that none of the information treatments change the amount of time participants invest in watching a video on how they can "*protect themselves from the virus*." At the same time, there are important effects on how much time people invest in watching a video about protecting the elderly. Learning only about your risk, which people tend to overestimate, reduces the share of participants watching a video titled *"how to protect old people"* by 9 p.p. (28%). Importantly, once we add information about the risk that older people face, this reduction in preventative effort disappears. The attenuation effect of information about the elderly seems to be driven by our invocation of specific, perhaps ironically more available, examples. Overall, we do not find reductions in preventive measures among respondents learning (on average) that their risk is lower than they thought, which we might have expected to be the rational response. Once we add information of a more vulnerable group whose risk participants tend to underestimate, they donate more to the CDC and invest more time in learning how to protect others.

While we find these results informative to the policy question we posed above, the results are smaller in magnitude and noisier than we might have expected. To understand the mechanisms behind the overall muted effects of our debiasing treatments, we conduct a third experiment with a new group of participants. The results of this experiment suggest that providing information about actual risk only leads to a partial updating of risk beliefs. Most participants lower their risk perceptions but still tend to overestimate the CMR despite saying the information we provide them is both relevant and credible. In addition, participants continue to extrapolate linearly away from the information they are given and this means they will continue to misestimate the risks of dissimilar others. Lastly, we find evidence that the "just-noticeable differences" of many people are rather large when it comes to risk perceptions (O'neill, 1977; Wilde, 1982). Specifically, many participants describe the risk as "low," while routinely overestimating it, and then go on to report that the objective (debiased) risk is not much different than their own initial perception, despite the actual risk often being an order of magnitude smaller.

We contribute to a large literature on debiasing risk perceptions through information provision (Lichtenstein et al., 1978; Weinstein and Klein, 1995; Robb et al., 2008; Abel et al., 2020; Akesson et al., 2020; Eil and Rao, 2011; Garrett et al., 2018). The overall effects of these interventions are mixed and our third experiment sheds light on the underlying mechanisms that drive the impacts of information debiasing interventions. Our results, for instance, suggest that people's perception of these treatments often differ from the (objective) informational content, and that it is their perceptions that appear to drive behavioral responses to treatment. It is thus important to take into account emotional reactions and subjective categorizing of new information when designing information interventions aimed at changing behavior.

Our study also adds to a nascent but exploding literature on COVID-19. Other studies have documented incorrect perceptions of infection rates (Fetzer et al., 2020; Akesson et al., 2020; Sjåstad and Van Bavel, 2020) and mortality rates (Akesson et al., 2020). In contrast to these studies, we elicited the entire age-risk gradient, which is important because misperceptions are not uniform across this gradient. While other studies have linked risk perceptions to the cognitive reflection test (Pennycook et al., 2020; Erceg et al., 2020; Frederick, 2005; Stanley et al., 2020) or media exposure (Simonov et al., 2020; Pennycook et al., 2020; Bursztyn et al., 2020), to our knowledge, we are the first to provide experimental evidence on how behavioral heuristics contribute to the misperceptions of the COVID-19 risks.

Similar to other studies, we explore the link between COVID-19 risk perceptions and pro-social behavior (Branas-Garza et al., 2020; Campos-Mercade et al., 2020).[3] However, we go beyond self-reported behaviors and use incentivized outcomes (donations and time investment). This turns out to have important implications for the conclusions we draw about the effectiveness of our debiasing intervention and echoes Falco and Zaccagni (2020) who find that the effects on self-reported behavior of a COVID-19 information intervention in Denmark do not necessarily predict actual behavioral changes.

Last, we add to the large literature on public good provision and prosocial behavior. A seminal model by Schwartz (1977) highlights that awareness of benefits is an important determinant of prosocial behavior. Our results show that people underestimate the risk to the most vulnerable population, suggesting they also underestimate the externalities they impose on others by not acting prosocially. The debiasing intervention finds that focusing people's attention on their own risks and benefits can reduce pro-social behavior and thus diminish the provision of public goods. However, adding information on the benefits of prosocial behavior by highlighting the high risk that others face can counter these adverse effects. These results are in line with studies from other domains showing that increasing awareness of externalities increases contributions to public goods (Dhont et al., 2012).

[3] One notable exception in this literature that goes beyond individual behavior is Bursztyn et al. (2020) who find that misconception induced by media sources leads to an increase in COVID-19 cases and deaths.

2 Study Design

2.1 Recruitment and Data collection

Because we wanted both more demographic and geographic variation than would be easily achieved in the lab, we recruited participants in the United States from Amazon's Mechanical Turk (MTurk) online platform. MTurk was originally created as an online workplace where people are paid to complete short Human Intelligence Tasks (HITs), but it has also become a popular tool for the recruitment of experimental participants. MTurk has become popular both because it provides convenient access to a more diverse subject pool and because studies suggest it can generate data that is at least as reliable as more traditional methods (Buhrmester et al., 2016; Merz et al., 2020). To improve data quality further, we exclude workers with HIT approval ratings below 96% and limited previous experience (i.e., less than 100 completed HITs).

As outlined above, we conducted several experiments with data collected in two waves. Institutional Review Board (IRB) approval was obtained from Middlebury College and the trial was registered at the AEA Social Science Registry (AEARCTR-0005579). For the first wave of data collection, a sample of 928 participants was recruited from March 19-22, 2020. The HIT title read "*Participate in a short health survey.*" Participants received \$0.72 base compensation plus an (unannounced) bonus of \$0.50. The task took, on average, seven minutes, resulting in a average hourly wage of \$10.45. In this first wave of data collection, we collected the perceptions of COVID-19 mortality rates used in Section 3. We then randomly assigned participants to receive information about actual (i.e., objective) mortality rates or a control group and collected data on prosocial behavior (details to follow in description of Experiment 2 in Section 5). From May 27-29 2020, we recruited our second wave of 387 participants using the same selection criteria and pay as the first wave to examine the underlying mechanisms behind risk perceptions (Experiment 1 in Section 3.2) and assess the efficacy of the debiasing treatment (Experiment 3 in Section 5).

2.2 Sample Characteristics

In the second column of Table 1, we report the characteristics of the participants in the first wave of data collection. Considering their demographics, these participants are 37.5 years old, on average, 38% are female and almost half have completed an undergraduate education. While MTurk samples are not always nationally representative, it is noteworthy that we match closely on age (the U.S. average is 38.2), we have broad geographic representation, with participants from 47 states and the District of Columbia, and we gathered substantial variation in political attitudes: 54% of respondents say they are liberal, 27% conservative and 19% moderate. Lastly, we find that, on average, participants answered half of the cognitive reflection task questions correctly, a success rate similar to random picnickers along the Charles River and Carnegie Mellon students (Frederick, 2005).

Table 1: Baseline Characteristics by Treatment (Wave 1, Experiment 2)

	N	Sample	Control	T1 (Own)	T2 (Gen)	T3 (Close)	T2+T3
Age	928	37.47	38.04	36.83	37.71	37.28	37.5
Female	928	.38	.4	.42	.35	.36	.36
4 yr college	928	.47	.44	.52*	.46	.46	.46
Cognitive Reflection	928	.56	.57	.56	.53	.57	.55
Liberal	928	.54	.5	.63**	.52	.53	.52
Moderate	928	.19	.19	.14	.2	.2	.2
Median Income (Zip)	928	65156	63804	65784	66637	64439	65528
Pop Density (Zip)	928	4929	4673	5300	5073	4670	4870
Deaths (Zip)	928	2.63	2.46	2.62	2.05	3.39	2.72
Corona cases (Zip)	928	203	217	167	173	255	214
Worried Corona	928	.33	.3	.35	.33	.33	.33
Corona Preventable	928	.18	.17	.16	.2	.2	.2
Mortality: Own	927	3.42	3.26	3.44	3.51	3.47	3.49
Mortality: Under 40	928	1.38	1.24	1.4	1.44	1.42	1.43
Mortality: 40-49	927	2.51	2.37	2.51	2.62	2.53	2.58
Mortality: 50-59	928	3.88	3.77	3.81	3.95	4.02	3.98
Mortality: 60-69	928	5.87	5.74	5.99	5.64	6.1	5.87
Mortality: 70-79	928	8.39	8.55	8.04	7.69	9.26	8.48
Mortality: Over 80	928	11.81	11.85	11.23	11.08	13.09	12.09

Notes: *Cognitive Reflection* measures the share of correct answers on the CRT. *Worried Corona* reports the share who believe they will contract the virus and *Prevent Corona* reports the share who believe they cannot protect themselves from the virus. *Mortality* indicates the perceived mortality risk for different age groups (winsorized at the 10% level). Column *Sample* reports average values for the full sample. The remaining columns report average values for the randomly assigned groups. Significance is reported for a test of equal means between the control and respective treatment groups. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1 also reports characteristics of the location where our participants reside. The zip code level median income is $65,156 compared to the national average of $59,000. At the time of the survey, the average number COVID-19 cases in participants' states was 507.9 and the average number of deaths was 7.1 compared to a national average of 479.8 cases and 6.7 deaths per state.[4] Considering how concerned our participants were with the COVID health risk, a third believed they would catch the virus and 18% thought they would be unable to protect themselves from the virus even if they took protective measures.

[4]Our average is a bit higher because the only three states we have no data from are ones with low CMRs: SD, ND, VT (New York Times, 2020).

3 COVID-19 Risk Perceptions

In this section, we first provide descriptive statics on people's risk perceptions and how these vary by demographic characteristics and factors related to respondents' exposure to COVID-19. We then provide evidence from our first experiment on the role of the availability heuristic on people's risk perceptions.

3.1 Subjective Risk Perceptions

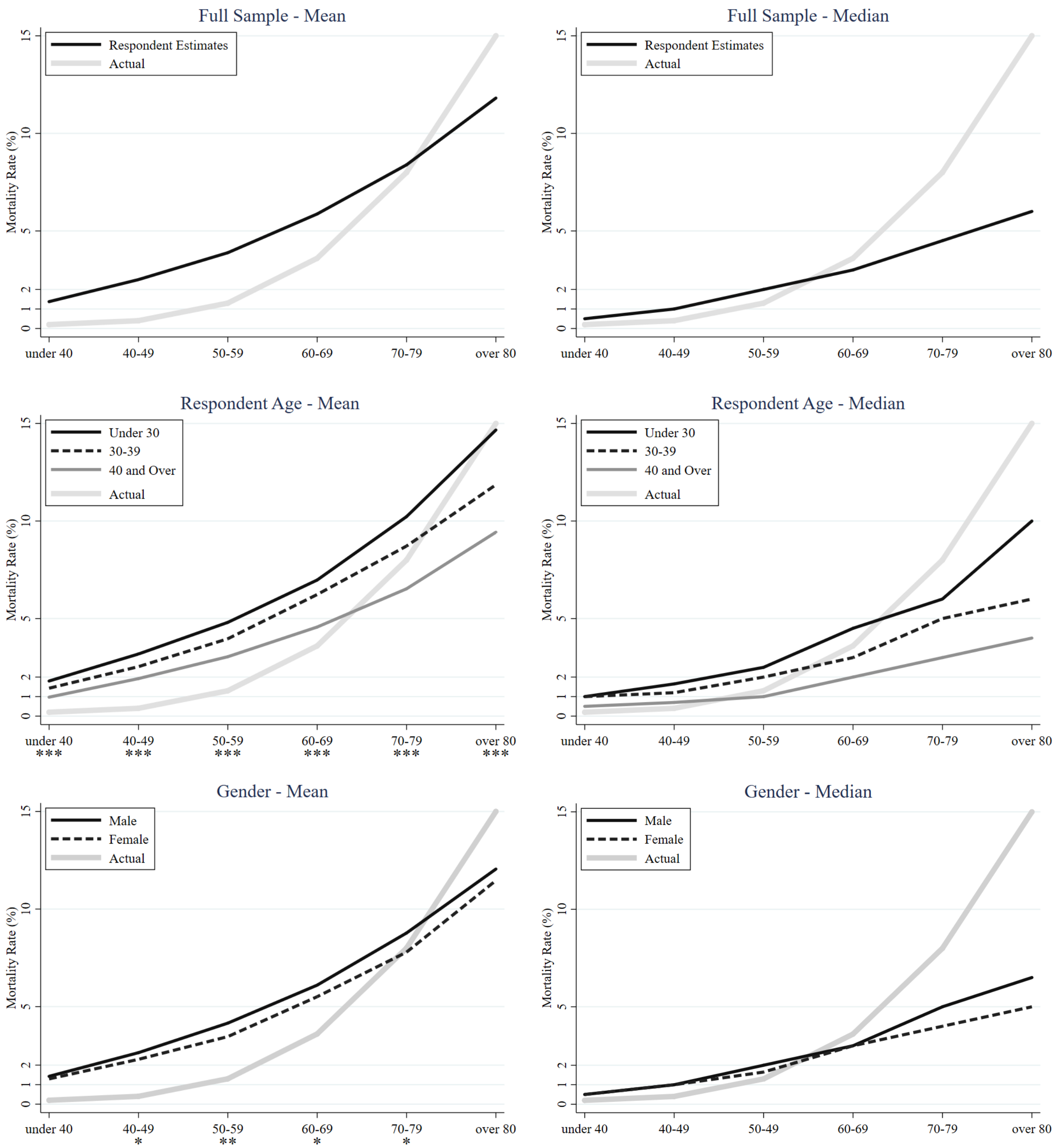
We begin the discussion of our results by describing demographic and behavioral patterns we find in the subjective risk perception data. We elicit a participant's beliefs about their own mortality risk of COVID-19 by asking them to respond to the following scenario: *"Suppose 1,000 people of your age are infected with the coronavirus, how many do you think will die from the virus?"* Further, to ensure that respondents do not confuse risk perceptions presented as frequencies from those described as percentages, we present the reported mortality rate as a percentage and give the participant the opportunity to correct their estimate (Appendix Figure A1). We then use the same method to elicit participant risk perceptions for six different age groups: Below 40, 40-49, 50-59, 60-69, 70-79, and over 80.

Figure 1 shows age-specific risk perceptions for different subgroups of our participants. The top panel shows average (left) and median (right) risk perceptions. As one can see on the left of the top panel, people on average substantially overestimate the risk for people up to the age of 70. This overestimation is particularly severe for the risk of younger people, where participants are off by an order of magnitude. By contrast, people are on average correct or slightly underestimate the mortality risk of older people. This masks substantial variation in risk perceptions. The top right panel shows that the median person's perception of the risk that people in their 70s and over 80 face is less than half of the actual rate. Another thing to notice about participant risk perceptions is that people tend to extrapolate linearly while the true age gradient is exponential. It is not clear if our respondents fail to understand that the risk profile is exponential in age or whether they know it is exponential and are unable to estimate risks exponentially away from their own circumstances.[5] Overall, this leads to situation where our mostly young participants overestimate the risk of people in their own age groups, but underestimate the risk that the elderly face.

The middle panels of Figure 1 show clear risk perception differences by respondent age

[5] Many studies have found instances of exponential growth bias which occurs when people treat exponential functions as strictly linear (Stango and Zinman, 2009; Levy and Tasoff, 2016). While we can reject that respondents estimate a linear risk gradient from the youngest to oldest ages, respondents' risk profile is statistically more linear than the actual risk gradient. The process of adjusting too little away from one's perspective (or another anchor) is well-documented (Tversky and Kahneman, 1974; Epley et al., 2004; Tamir and Mitchell, 2010).

Figure 1: Estimates of Mortality Risk Age Gradient by Respondent Characteristics



Notes: Estimates of means are based on risk perceptions from Wave 1 wizorized at a 10% level. Asterisks under x-axis labels in the responding age panel indicate p-value of Jonckeree Terpstra tests of rank ordering of risk perception by the relevant subgroup categorization. *** p<0.01, ** p<0.05, * p<0.1 Appendix Table A1 reports tests of risk ranking for each category of each subgroup for all age-risk levels. In the Gender panel, which only compares a binary relationship, the asterisks report the p-value of one-sided test that risk estimates are higher for male and female respondents.

(divided into terciles). Both on average, and at the median, the younger the respondents, the larger they perceive the risks to be for everyone. Asterisks under the age-group axis labels in Figure 1 indicate p-values of Jonckeree Terpstra tests of rank ordering of risk perceptions by the relevant subgroup categorization.[6] One explanation is that younger people are less likely to be informed about COVID-19 and thus are more likely to rely on heuristics in assessing risks.[7] It is noteworthy that the group with the lowest personal risk (i.e., those under 30) is also the group with the highest risk perceptions. This fact raises concerns about the social implications of debiasing the young, the experiment we discuss in Section 4. Will the young, if they adopt the correct, and much lower risk estimate for themselves (and continue to extrapolate linearly), pull back too much on preventive measures because they think the real threat is much lower for themselves and others?

The bottom panels of Figure 1 show risk perceptions by respondents' gender. In contrast to other studies (Akesson et al., 2020; Brody, 1984; Gwartney-Gibbs and Lach, 1991; Savage, 1993; Finucane et al., 2000), we find that men tend to have higher risk perceptions, though these gender differences are only significant for certain age groups. One possible explanation why these perceptions seem "flipped" is that the raw COVID-19 mortality rate, not conditioned on smoking, for example, was higher for men and reported in the media (Rabin, 2020).

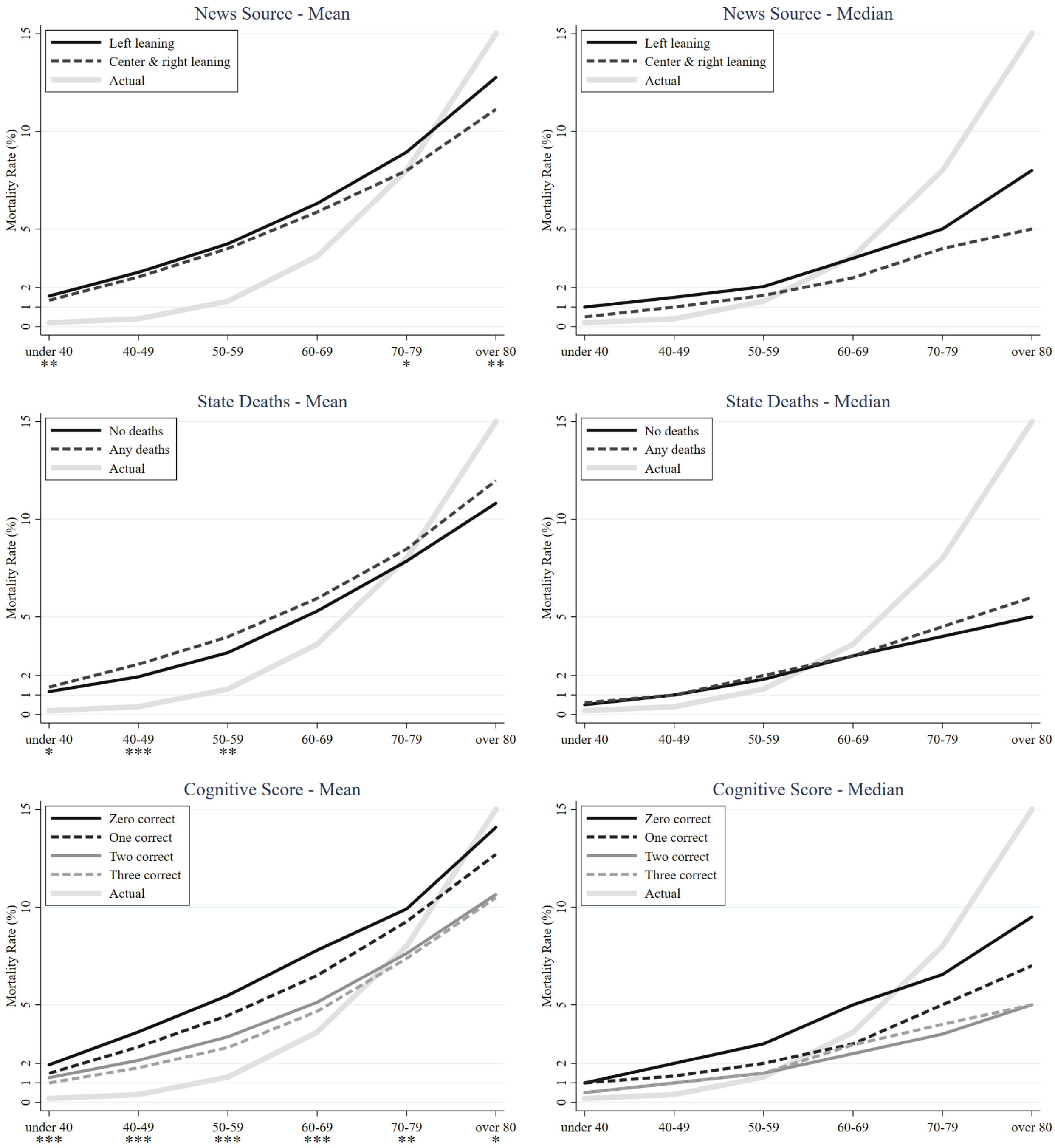
Our results on COVID-19 risk perceptions are very similar to Lichtenstein et al. (1978), which forms a cornerstone in the behavioral assessment of subjective health risk perceptions. People dramatically overestimate personal risks that are top of mind (i.e., available), perhaps because they are novel, vivid or discussed frequently in the media, and underestimate those risks that are harder to imagine or relate to, in this case the risk to people much older (Tversky and Kahneman, 1973). To further examine the role of the availability heuristic in forming subjective risk perceptions, we next compare how these beliefs vary by media exposure, whether respondents are potentially exposed to COVID-19 fatalities and a key measure of the propensity to think intuitively (i.e., rely on heuristics), instead of analytically.

In the top panel of Figure 2, we see that the political leaning of the news source that participants are most likely to consult predicts their assessment of the mortality risk of COVID-19. There is already plenty of evidence that the conservative media in the U.S. downplays the risk of the virus and that the liberal media might overestimate it (Simonov et al., 2020; Pennycook et al., 2020). We see the "availability" effect of viewing these different news outlets using both the average perceptions in the left panel and even more clearly at the medians on the right. Here the assessment of liberal media viewers (i.e., those whose

[6] For example, the p-value for the test that under-40 risk perceptions across respondents of different ages are equal versus an alternative that risks assessments decrease as respondent age increases has a p-value $<$ 0.01. Further, we explicitly test that risk estimates are higher for respondents under 30 than those 30-39 and higher for 30-39 year olds than for those 40 and over at each age-risk category. The results of these tests are reported in Appendix Table A1.

[7] Abel and Brown (2020) find that younger people report following the COVID-19 crisis less closely.

Figure 2: Estimates of Mortality Risk by Availability Features



Notes: Estimates of means are based on risk perceptions from Wave 1 wizorized at a 10% level. Asterisks under x-axis labels in the Congnitive Score panel indicate p-value of Jonckeree Terpstra tests of rank ordering of risk perception decreasing from zero to three correct answers on the CRT. Appendix Table A1 reports tests of risk ranking for each category of each subgroup for all age-risk levels and also provides statistical tests for comparisons at the median. In the New Source and State Deaths panels, which compare binary relationships, the asterisks report the p-value of one-sided test that risk estimates are higher for respondents who watch liberal media compared to right or center or that respondents from states with any deaths estimate higher risk than for those with no deaths. *** p<0.01, ** p<0.05, * p<0.1

news comes from MSNBC, CNN or PBS) is always above that of right leaning (FOX News) or center leaning media viewers (Bloomberg and ABC News) (Pew Research Center, 2014). The differences at the median are highly statistically significant for all age-risk categories (see Appendix Table A1).

The middle panels of Figure 2, show how risk perceptions depend on whether there were any COVID19-related deaths in a respondents' zip code. In principle, viral risks are more available if they happen more frequently where you live, regardless of their relative frequency, and this is what we see. Participants from zip codes that report any deaths have higher risk perceptions than those that do not.

Lastly, the bottom panel of Figure 2 indicates that there is a particularly strong monotonic relationship between one's score on the cognitive reflection test, which has recently been linked to the use of heuristics like availability (Toplak et al., 2011; Johnson et al., 2014; Morsanyi et al., 2014) and one's risk perception. Respondents who get none of the CRT questions right (and almost uniformly give the intuitive but wrong answers), have risk perceptions that are much greater than those respondents who get all three questions right. Further, those answering one question correctly have lower risk perceptions than those answering none correct and participants answering two correctly have even lower perceptions. Importantly, the better one does answering these questions, designed to sort between "intuitive" and "analytical" thinkers, the closer one is to having the correct risk perceptions.

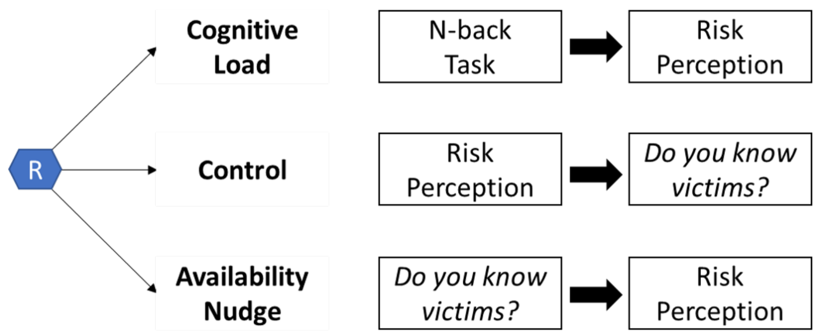
Taken as a whole, the evidence presented in Figure 2, indicates that the availability heuristic is a strong candidate to explain the biased risk perceptions we find. While it is reassuring that results are consistent across several measures, the evidence in this figure is still correlational. In the next section, we present experimental evidence from our second wave of data collection that further corroborates this conclusion.

3.2 Risk Perception Mechanisms

3.2.1 Experimental Design

We conducted an experiment in May 2020 with 287 participants to examined the links between COVID-19 risk misperceptions and the availability heuristic by experimentally manipulating two pathways. First, if the mere volume of news and bandwidth attributed to the virus occupies a significant portion of one's working memory, people should be more reliant on heuristics, as hinted at with our CRT results in Figure 2. To directly test this mechanism, we exogenously imposed a cognitive load on one's working memory and test if doing so increases risk perceptions in a similar way. Second, if the ease with which people can retrieve virus mortalities from their working memory determines the magnitude of their risk perceptions, like the people in Figure 2 who have witnessed deaths in their zip codes or have heard of more deaths on their devices, then exogenously enhancing this recall should make instances more available and increase one's risk perceptions.

Figure 3: Experiment 1 Design: Availability Mechanisms

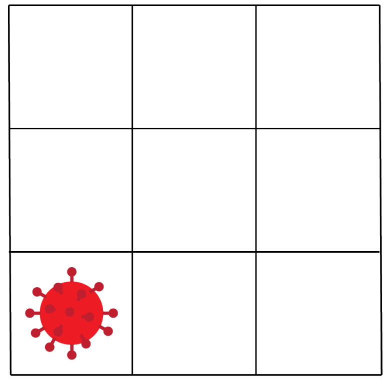


In addition to a new control group that replicated our first wave (i.e., March, 2020) design, this second set of participants was randomized into two treatments (see Figure 3).[8] In the cognitive load treatment, while participants were filling in their risk perceptions, they were incentivized to watch and pay attention to the three-by-three grid in Figure 4 and keep track of the position of the COVID virus which moved randomly every two seconds. If when the virus stopped moving, a time strategically titrated to be just about when most people submitted their risk perceptions in the first wave of data collection, and a person could remember the penultimate position of the germ in the grid (a position they had to remember until the next screen appeared and they could type it in) they received a bonus payment of 15 cents (20% of the base compensation). In other words, our procedure combined both aspects of standard "N-back" cognitive load manipulations - we loaded up one's working memory and we divided their attention (Owen et al., 2005). At the end of the cognitive load treatment, we did a manipulation check with participants using the customary NASA-TLX subjective assessment of workload (Hart and Staveland, 1988). Specifically, these participants were asked how mentally demanding the task was and how hard they had to work to accomplish their level of performance.

In the availability nudge treatment, we first asked participants, *do you know of someone in your community or social network who has died of COVID-19?* The purpose was to bring any mortalities to the top of one's mind before they reported their risk perceptions. We also asked whether the person they were imagining was a relative, friend or community member. Control group participants were asked the same availability nudge questions *after* we elicited their risk perceptions to allow testing whether treatment effects differ by whether people knew victims. The shares of people knowing a person dying of COVID-19 were 10.7% for friends, 14.3% for relatives, and 13.3% for community members.

[8] Appendix Table A2 shows that only one out of 32 pairwise comparison of baseline characteristics yields significant differences at the 10% level suggesting that Wave 2 randomization was successful.

Figure 4: Cognitive Load Task (Experiment 1)



Notes: In this N-back task, participants were paid a bonus if they could remember the penultimate position of the virus when it stopped moving randomly about the grid.

3.2.2 Replication

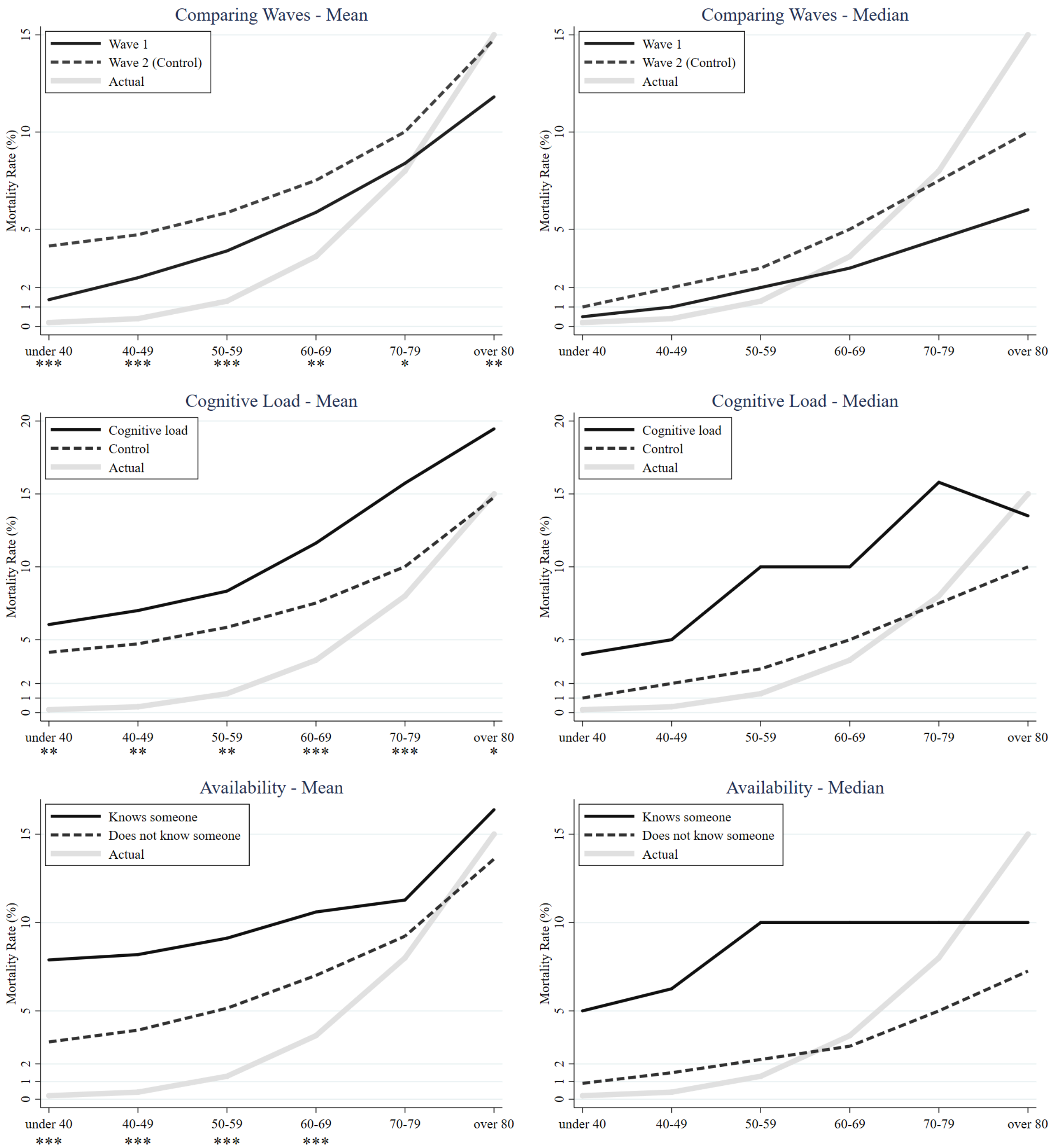
The obvious first question concerning our second wave of data collection is whether or not we replicated our earlier results. At the top of Figure 5, we see that our second wave yields a very similar age-risk gradient as in Figure 1. Participants tend to vastly overestimate their own risk of COVID-19 and continue to underestimate the risk to older people (both on average and at the median). In fact, we find that risk perceptions were even higher in May than they were in March particularly for the risks at younger ages. Even this increase is consistent with an availability bias in subjective risk perceptions. Not only had the media attention continued to focus on the virus in the intervening months, the number of fatalities increased making them easier to envision and recall.

3.2.3 Cognitive Load

According to their self-reports, most participants in the cognitive load treatment found the 2-back task burdensome. On a 10-point Likert scale, participants rated the mental demand of the task at 6.2, on average, and they felt they had to work with an average intensity of 6.9 to succeed at the task. These subjective assessments are borne out in the task results. Just 21% of people could accurately report the penultimate position of the germ, though another third were off by just one position.

How did the imposition of the load affect risk perceptions compared to the control condition? The middle panel of Figure 5, indicates that loading up a participant's working memory makes their subjective risk perceptions even more dire. The increase in risk per-

Figure 5: Estimates of Mortality Risk Age Gradient by Arm (Experiment 1)



Notes: Estimates of means are based on risk perceptions wizorized at a 10% level. In the Comparing Waves panel, asterisks under x-axis labels indicate p-values of one-sided tests that risk estimates are higher for in Wave 2 than in Wave 1. *** p<0.01, ** p<0.05, * p<0.1. The Cognitive Load and Availability panels present data from Wave 2 only. In the Availability panel, we pool together respondents from the Control and Availability arms.

ceptions is large and significant at every point along the age-risk gradient (Appendix Table A3).

While an extensive literature shows that imposing a cognitive load increases reliance on heuristics, it was not clear (ex-ante) that risk perceptions would be determined by the availability heuristic, instead of being affected by one of the many other heuristics. Importantly, we find that the load caused an increase in risk perceptions, an effect consistent with the availability heuristic, but not with the obvious other candidate heuristics and biases. Optimism bias, best-case heuristic, overconfidence and the illusion of control, for example, would all predict that risk perceptions decrease (DeJoy, 1989; Bränström et al., 2006; Simon et al., 2000; Houghton et al., 2000; Broihanne et al., 2014; Sjåstad and Van Bavel, 2020). Hence, the cognitive load intervention is also consistent with our participants reliance on availability and not the other common biases in risk perceptions. Next, we discuss the results of a more narrow test of the availability heuristics from the other arm of Experiment 2.

3.2.4 Availability Nudge

The bottom panel of Figure 5 shows that the risk perceptions of respondents who knew someone who had died of COVID-19 are on average about twice as large compared to those of people who did not know a victim. Table 2 further shows that these risk perception differences are driven by respondents who lost a friend or family member to COVID-19. These findings are in line with predictions of the availability heuristic, since it is easier to conger an image in one's working memory of a friend or relative who has succumbed to the virus than it is to imagine a more anonymous community member. These results are also consistent with Nisbett and Ross (1980) and Keller et al. (2006) who find that more salient and emotional events induce availability and affect risk perceptions more.

Table 2 also indicates that the treatment effect of nudging people to recall this person before estimating risk, (i.e., bringing a COVID-19-related death to the top of a respondents' mind) is positive and sizeable for most age groups, but not statistically significant. It is also noteworthy that the impact of whether a person knew of a COVID-19 victim does not depend on whether they were nudged to recall this person before estimating risk (Appendix Table A4). One explanation is that knowing a victim is already a very salient event, so the availability nudge may not have had an additional effect. Likewise, the initial survey questions about COVID-19 may have already increased the salience of the risks for those knowing victims. In other words, what is most important in driving risk perception is knowing a victim, not whether we nudged respondents to recall this.

Overall, we interpret the evidence presented in this section as support for the importance of the availability heuristic, the basis for the social dilemma that can arise when inflated risk perceptions are debiased. Congruous with availability, our participants subjected to a cognitive load also dramatically overestimate the risk for all but the very old and those other

Table 2: Treatment Effects: Availability and Risk Perceptions (Experiment 1)

	Risk Perception						
	Own Risk	Under 40	40-49	50-59	60-69	70-79	Over 80
Availability Nudge	21.44	12.60	11.40	11.99	12.25	-3.423	-11.04
	(24.64)	(8.807)	(8.988)	(9.628)	(11.73)	(14.00)	(22.70)
Friend died	168.8***	76.11***	62.49***	65.92***	69.31***	52.24*	52.55
	(49.25)	(16.85)	(16.65)	(18.07)	(20.73)	(21.22)	(37.97)
Relative died	151.8***	37.36*	35.30*	21.71	24.37	4.064	-2.680
	(44.94)	(15.83)	(15.10)	(13.28)	(16.07)	(17.99)	(30.00)
Com. member died	4.325	-7.307	5.119	6.126	11.66	3.937	34.97
	(40.64)	(12.30)	(12.45)	(14.91)	(16.34)	(17.38)	(36.89)
R^2	.171	.171	.129	.106	.084	.029	.018
Observations	195	195	195	195	195	195	195
Sample Mean	104.54	42.76	49.66	61.54	82.95	109.85	155.45
Std Dev	169.15	63.60	63.741	68.23	83.98	105.275	163.08

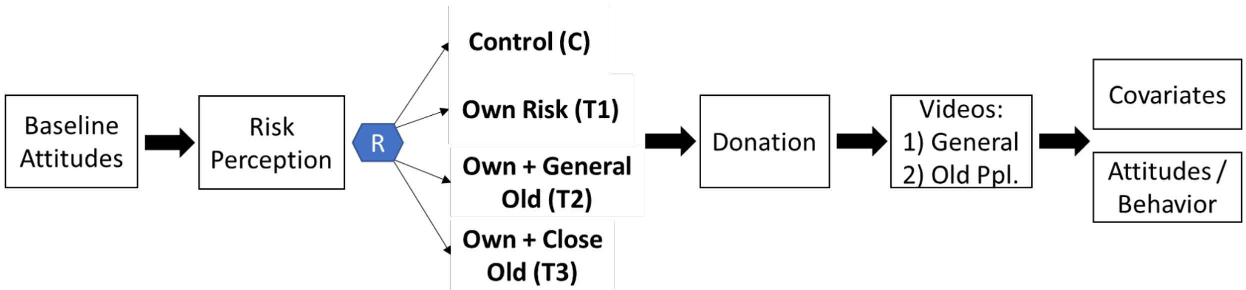
Notes: Dependent variables measure the mortality risk perceptions (out of 1,000 infected people) for different age groups. *Availability Treat* measures whether a participant was asked about knowing a victim before risk perceptions were elicited. The other variables measure whether participants know of a friend, relative, or community member that died of COVID-19, repectively. All estimations are OLS. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

participants who are nudged do appear to have higher perceptions as well. Further, while knowing a victim is endogenous, it is noteworthy that the coefficients in Table 2, are robust to controlling for characteristics like age, gender, and education.

4 Mortality Risk Debiasing Experiment

Now that we have presented our findings on the risk perception patterns we found in both waves of data collection and demonstrated that these patterns are consistent with the availability heuristic, we proceed by describing the experiment we ran as part of the first wave of data collection. This experiment asks whether debiasing risk misperceptions affects prosocial behavior. In this case, the underlying hypothesis (from the rational actor point of view) is that since objectively most participants will receive good news, in that their own actual risk is much lower than anticipated, they will pull back on costly efforts to protect themselves and others. Simply put, debiasing may lead to a social dilemma.

Figure 6: Experiment 2 Design: Debiasing Risk Perceptions



4.1 Experimental Design

Participants were randomized into one of three treatment arms or a control group (see Figure 6). The comparison of characteristics by treatment arm in Table 1 suggests that randomization was successful.[9] For all of the information treatments, we used the most current mortality data available in March 2020, published in the Lancet (Verity et al., 2020).[10] The first treatment group (T1) received information about the mortality risk of people in their age group, which in most cases differed substantially from participants' prior beliefs. The information provided states, for example, *You thought that 15 out of 1,000 infected people your age will die. A recent study found that the mortality rate of people below the age of 40 is 0.2%. This means that 2 out of 1,000 infected people in this age range die from the disease.* A supporting graph further visualized this information (Figure 7, left panel).

Treatment group 2 (T2) received two bits of debiasing information. In addition to seeing the same, personalized, information as in T1, participants in T2 received information about mortality rates of elderly people. An example of this information reads, *You thought that 50 out of 1,000 infected people over the age of 80 will die. A recent study found that the mortality rate of people over the age of 80 is 14.8%. This means that 148 out of 1,000 infected people in this age range die from the disease.* Treatment group 3 (T3) received the same, personalized, information as in T1 as well as as mortality rates of a known elderly person in an attempt to make the risk to old people more available than in T2.[11] Specifically, we first asked participants for the name, relationship to and age of a person older than 70 that they "are close to." The graphic on the right side of Figure 7 illustrates a concrete example

[9]Of the 57 possible pairwise comparisons of characteristics between the treatments in our experiment and the control, we find only two significant differences at the 10% level, suggesting that randomization was successful (Table 1). We also analyze how effects differ between treatment arms. Of 108 pairwise tests of equal means, four differences are significant at the 5% level and three are significant at the 10% level.

[10]This study used data from China. Some projections of mortality rates for the U.S. were lower at the time, which would lead to even larger overestimations of risks than those reported in Sections 3 and 4. Importantly, the results discussed below show that a large majority of participants thought that the information was both credible and relevant for them.

[11]To make the information in T2 and T3 as similar as possible, T2 provides information about either people in their 70s or over 80 based on the share of people specified in T3

Figure 7: Debiasing Information Treatments (Experiment 2)

Notes: Participants in the debiasing experiment were shown some combination of these two graphics, individualized to their responses. T1 participants saw just the information of the left, T2 participants saw both bits of information, minus the personalized first sentence at the top of the right panel and T3 participants saw it all.

of this treatment. The information treatment states, *"Your Grandmother Jane Mead's age is 82. You thought that 60 out of 1,000 infected people over the age of 80 will die. A recent study found that the mortality rate of people over the age of 80 is 14.8%. This means that 148 out of 1,000 infected people in this age range die from the disease."*

For all participants, we collect three types of outcome measures: the time they invest in learning about preventative measures, how much of their compensation they are willing to donate to the CDC's COVID-19 Emergency Fund, and as a link to recent related studies, self-reported prosocial behavior and attitudes. Details about these outcome measures and results are discussed below.

4.2 Debiasing Experimental Results

To estimate the treatment effects of providing debiasing information we estimate the following two equations, as specified in our pre-analysis plan, using OLS and robust standard errors.

$$y_i = \beta_0 + \beta_1 Town_i + \beta_2 Told_i + X_i' \beta + \epsilon_i \tag{1}$$

$$y_i = \beta_0 + \gamma_1 T1_i + \gamma_2 T2_i + \gamma_3 T3_i + X_i' \beta + \epsilon_i \tag{2}$$

Here y_i measures outcome y for participant i, $T1$ through $T3$ are indicators for the information treatments, $Town$ is identical to $T1$, $Told$ is an indicator that aggregates $T2$ and $T3$ and X is a vector of controls from Table 1. The only difference between these two empirical models is that we pool the two types of information about the risk to older people. β_1 therefore estimates the effect of learning about your own mortality risk and β_2 estimates the effect of learning about the mortality risk of old people generally (in addition to your own risk). We also report results from specification (2), which estimates the effects of learning about the mortality risk of an unspecified old person (γ_2) and the risk to a close elderly person (γ_3) separately.

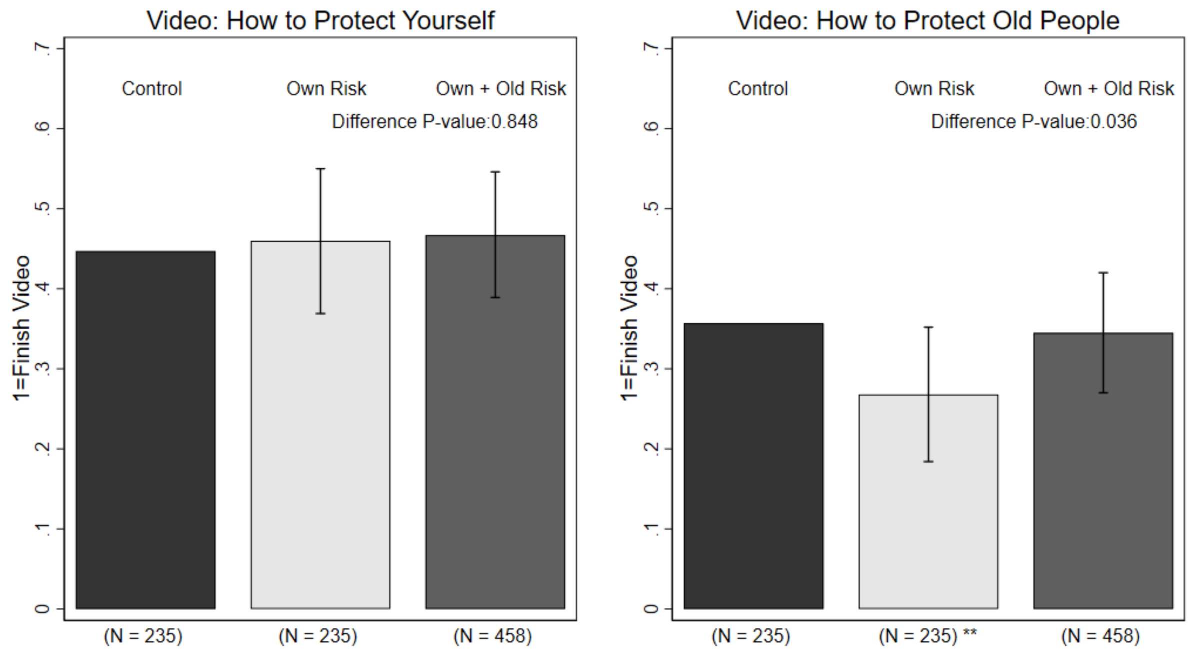
Being more specific about the prosocial outcomes we gathered in this experiment, we classified watching two informational videos produced by the CDC about how to protect yourself and how to protect older people as attention to preventative measures.[12] The first video about protecting yourself lasted 36 seconds and the second video about protecting older people was 145 seconds long. It is costly to invest time in this knowledge and, in this sense, these two outcomes are incentivized. The third outcome is also incentivized. Here we allow participants to donate up to 50 cents of their earnings to the emergency fund that the CDC established to help deal with the pandemic. Though un-incentivized, we also included a few self-reports of prosocial behavior to compare our results to those of others utilizing similar outcomes. These behaviors included cancelling a dinner party because of the virus, wearing a face mask to prevent the spread and paying higher taxes to make sick leave available to all workers.

Starting with the videos on how to stop the spread of the virus and protect older people, in Figure 8 we report the average fraction of participants who finished watching each video, by condition. Surprisingly, and suggesting that the social dilemma of debiasing may not be a dire as expected, we see on the left that there is absolutely no effect of learning about one's own actual risk or the actual risk to older people on whether participants finish the video about how to protect themselves. Regardless of the information provided, approximately 45% of participants finish this video. This conclusion is confirmed in Appendix Table A5, which presents estimates of the treatment effects with and without controls and combining T2 and T3 or estimating the effects separately.

At the same time, we see on the right side of Figure 8 that learning, for most, that you have overestimated your own mortality risk does present a problem for society because these people are 9 percentage points (pp) less likely to finish the "How to Protect Old People" video. This constitutes a 28% fall compared to the control. Thankfully, we see that adding the actual risk information for old people, which for many indicates that they have underestimated this risk, attenuates the drop in prosocial behavior. Here, the "Own+Old Risk" bar is of the same height as the control treatment. Assuming that the effects of information on "own"

[12]The first CDC video was titled, "How to protect yourself against COVID-19" and the second was titled, "What Older Adults Need to Know". The former video is not available online anymore, the latter video can be found on the CDC website and here.

Figure 8: Treatment Effects: Attention to Prevention (Experiment 2)



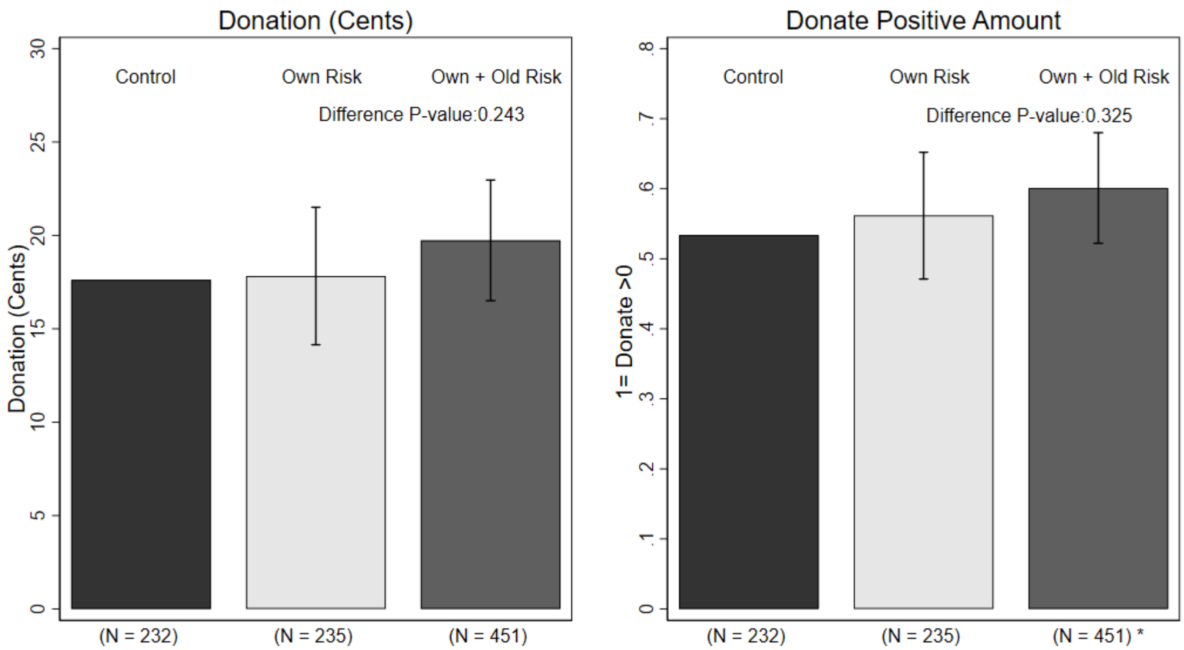
Notes: The graph shows the share watching the video on how to protect yourself (left graph) and on how to protect others (right graph) by treatment status. Significance levels for a test between treatment arms and control group are indicated below the bars next to the sample size. The p-value for the difference between T1 and the combined T2 and T3 are reported under "Difference p-value".

and "old" people are additive, these results suggest that learning about old people's risk increases interest in learning about how to protect these people by 7.9 pp (22%). Again, these results are confirmed in Appendix Table A5, where we also see that this positive effect is driven largely, as hypothesized by information about close (perhaps more available) old people. This information, in particular, increases the share watching the video by 10.4 pp compared to a 4.9 pp increase for the general information about old people. However, this difference between T2 and T3 is not statistically significant (p-value=0.21).

Overall, these results suggest that providing information about your own risk reduces investments in attention to protect others (i.e., does pose a social dilemma). However, additionally providing information about the risks that the most vulnerable group faces may offset these effects suggesting that information about older people has the potential to increase prosocial behavior.

After our participants watched as much of the two videos as they wanted, we informed them that they would receive a $0.50 bonus for completing the survey. Participants then had the chance to donate any amount of this bonus to the CDC Emergency Fund which, *will be used to respond to the public health threat posed by this virus. The funds go to additional*

Figure 9: Treatment Effects: Donation to CDC Pandemic Relief Fund (Experiment 2)



Notes: The graph shows the average amount people donate (left graph) and the share who donate a positive amount (right graph) by treatment status. Significance levels for a test between treatment arms and control group are indicated below the bars next to the sample size. The p-value for the difference between T1 and the combined T2 and T3 are reported under "Difference p-value".

support for personal protective equipment and critical response supplies, which may help to prevent the spread of the coronavirus. Appendix Figure A2 is a screen shot of the exact solicitation.

Figure 9 graphs the treatment effects of our debiasing information on giving to the CDC's Emergency Fund. As one can see, on neither the extensive (Donate Positive Amount) nor the intensive (Donation) margins, is there an effect of learning about your own actual risk on charitable giving. Like the null effect on watching the video to protect yourself in Figure 8, these results suggest that debiasing people with their own risk information does not produce negative externalities.

Again, like our video results, we find as shown in Figure 9, that debiasing participants with information about older people (the risk to whom they tend to underestimate) leads to a modest increases in donations. The average amount donated increases by 2.2 cents (11.2%, p-value=0.156) and the share donating a positive amount increases by 6.6 pp (15.8%, p-value=0.089). However, differences to the own risk group are not statistically significant. All the results described in Figure 9 are presented in more detail in Appendix Table A6.

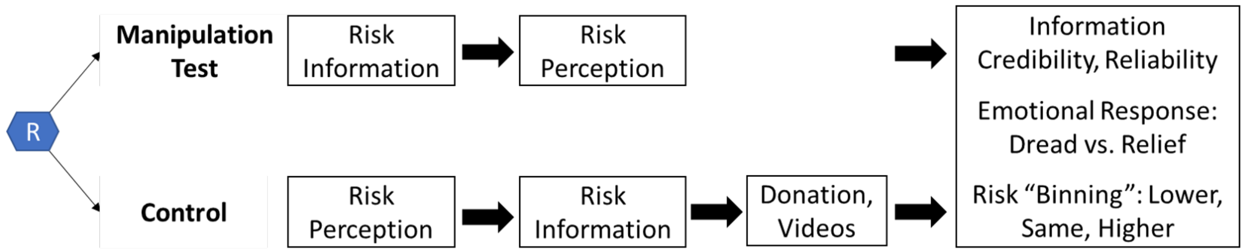
Lastly, considering the un-incentivized self-reports of prosocial behavior, we find either no influence of our debiasing information treatments or counter-intuitive results. A full analysis of the outcomes is presented in Appendix Table A7 but a summary is as follows. When participants reported how likely they were to cancel a dinner party because of the virus, being shown their own actual risk information had no significant effect but being shown the actual risk to older people, generic older ones in particular, they actually reduced their willingness to cancel dinner ($p < 0.05$). None of the information treatments had a significant effect on our participant's willingness to wear a face mask or pay an additional tax to provide sick leave for everyone.

Summing the impacts of our debiasing experiment on the incentivized outcomes, we find effects that depend on the type of externality generated. Our outcomes represent two types of prosocial acts: acts with both internal and external benefits (watching video 1), and acts that are purely external/altruistic (watching video 2 and making donations). When participants learn that their own risk is considerably lower than they thought (as occurs in most instances), it does not make them reduce prosocial acts that mostly benefit themselves (watching video 1) but it may lead them to reduce acts that are purely altruistic (watching video 2). This behavior is partially consistent with the predictions of the standard rational actor model and can lead to a social dilemma. In contrast, when participants receive news that the risk is higher for old people (as most people do), they tend to invest more in purely prosocial acts specifically targeted to benefit these more vulnerable people (watching video 2 and making donations).

5 Why is Debiasing Risk Perceptions (In)Effective?

Overall, debiasing had mixed effects on prosocial behavior. Providing information only about your own risk has either no effect or may even decrease prosocial behavior. Combining it with information about the risk of older people can mitigate some of these negative effects and may lead to a modest increase in charitable giving. This section discusses evidence from a third experiment we conducted with 193 participants as part of the Wave 2 data collection to understand the mechanisms that drive the impact of debiasing (or lack thereof) on prosocial behavior. The goal of this third experiment is two-fold. First, we assess whether participants found the information relevant and credible and determine whether the information treatments updated respondent beliefs (Section 5.1). Second, we use questions about subjective perceptions of the information to explore heterogeneous impacts of our debiasing treatment on prosocial behaviors (Section 5.2).

Figure 10: Experiment 3 Design: Manipulation Test



5.1 Information credibility/relevance and belief updating

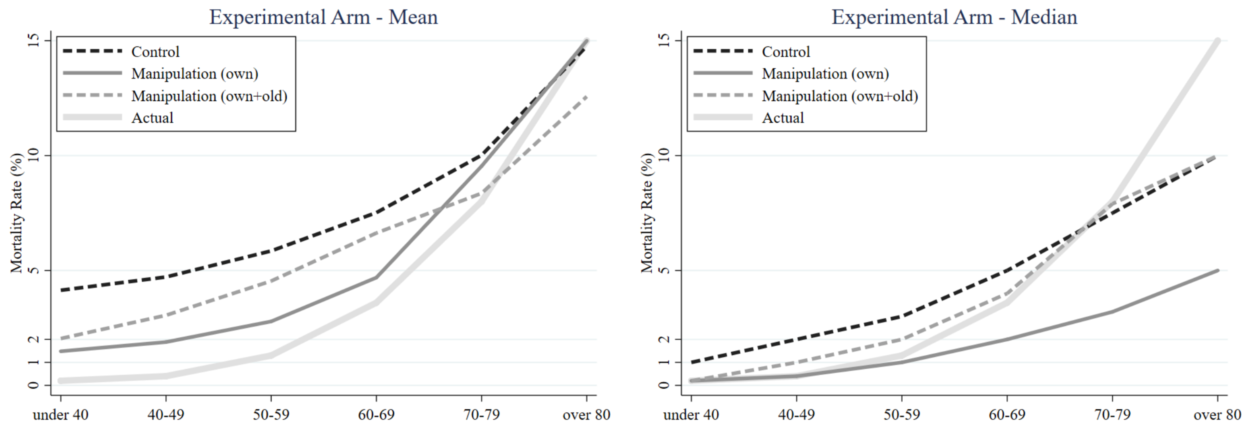
At the beginning of Experiment 3, we randomized whether risk information was provided before or after we elicited respondents' risk perceptions (see Figure 10). This allows us to test whether our information treatment changed respondents' mortality risk perceptions. We also included a set of questions at the end of the survey about the credibility and relevance of the mortality risk information we provided.

One hypothesis concerning the effectiveness of debiasing comes from the well-established fact that there is considerable suspicion with regard to medical information in parts of the population (Matthews et al., 2002; Hesse et al., 2005; Rains, 2007; Lipset and Schneider, 1987; Alsan and Wanamaker, 2018). With this in mind, we asked participants, *How credible do you find the information you were shown about the coronavirus mortality rate?* We find that only 6.8% of respondents state that they did not find the information credible.

Another factor that may have decreased the effectiveness of the treatment is a lack of perceived relevance of information, given that most of our participants are in their 20s and 30s. For example, Weinstein and Klein (1995) found that undergraduate students did not respond to information debiasing treatments about health risks. This is in line with other evidence showing that people are overly optimistic about their personal state of health (Weinstein, 1989). We thus asked participants, *How relevant is this information for a person of your health status and location of residence?* We find that only 7.8% state that the information was not relevant.

Given that the information we provided was perceived to be credible and relevant for most participants, it is unsurprising that respondents randomly assigned to receive the information treatment *before* reporting their risk perceptions provide significantly lower risk assessments, as seen in Figure 11. It is, however, noteworthy that the updating is incomplete, even for the risks at ages for which we provided specific information. Considering the under-40 risk, for example, those who received only this information before making their risk assessments estimated the mortality rate to be just under 2 percent which is significantly lower than the average estimate of 4 percent among those who were shown the information after; but 2 percent is still ten times larger than the actual rate of 0.2 percent for that age group.

Figure 11: Estimates of Mortality Risk Age Gradient, Manipulation Arm (Experiment 3)



Notes: Estimates of means are based on risk perceptions from Wave 2 wizorized at a 10% level.

Figure 11 also indicates that there is interesting heterogeneity in belief updating across those who were randomized to receive only the information about the risk for their age group (own) versus those who also learned about mortality risk for older individuals too (own+old). For many participants, it is not obvious that receiving more information gets their risk perceptions closer to the actual age-risk gradient. As the left panel indicates, those receiving just their own information come closest to the actual gradient, on average, and those who receive both bits of information tend to continue to overestimate the risks, on average. That said, at the median those who only receive information about their own risk lower their risk perception for the elderly, leading to an increased underestimation of the risk at older ages, again due to linear extrapolation. This may explain the negative effect of the own information treatment (T1) on the share of people watching the video about how to protect elderly people.

5.2 Subjective Categorization of Risk Information

Using the Experiment 2 sample, we first test whether the effects of the risk information on watching videos (Figure 8) and donations (Figure 9) vary by participants' initial misperceptions. We hypothesized that learning that COVID-19 mortality risks were higher than initially thought could nudge people to act more prosocially or conversely that learning that they were lower could reduce prosocial behavior. Contrary to this hypothesis, in Appendix Table A8, we find no evidence of those who initially underestimated the risk acting more prosocially after receiving the information. The effect of misperception is, in most cases, insignificant and for some outcomes even in the opposite direction (e.g., learning that one's own mortality risk is lower than expected, leads to larger donations (Appendix Table A8).[13] However, this categorization of misperception is based on our objective measure, as economists and behavioral researchers, of over- and underestimation. As such, we are concerned that the debiasing treatment, as measured by us the *researchers*, may not match the *respondents'* perception of the information treatment. In the end, it is the perceptions of the participants that should effect their prosocial behavior.

To understand participants' subjective perceptions of the information treatment, we collected data on how respondents categorize new information and to what extent the information triggers an emotional response. Specifically, we asked people in the manipulation control arm, *Did the information you were provided make you think the risk was: much LOWER than you originally thought, much HIGHER than you originally thought, or about the SAME as you originally thought.* Because mortality risk is an emotional topic, people's reaction to this new information may differ from that of a rational agent. To better explore this affective mechanism, we also asked respondents to rank how the information we provided made them feel on a scale from 0 (relief) to 10 (dread).

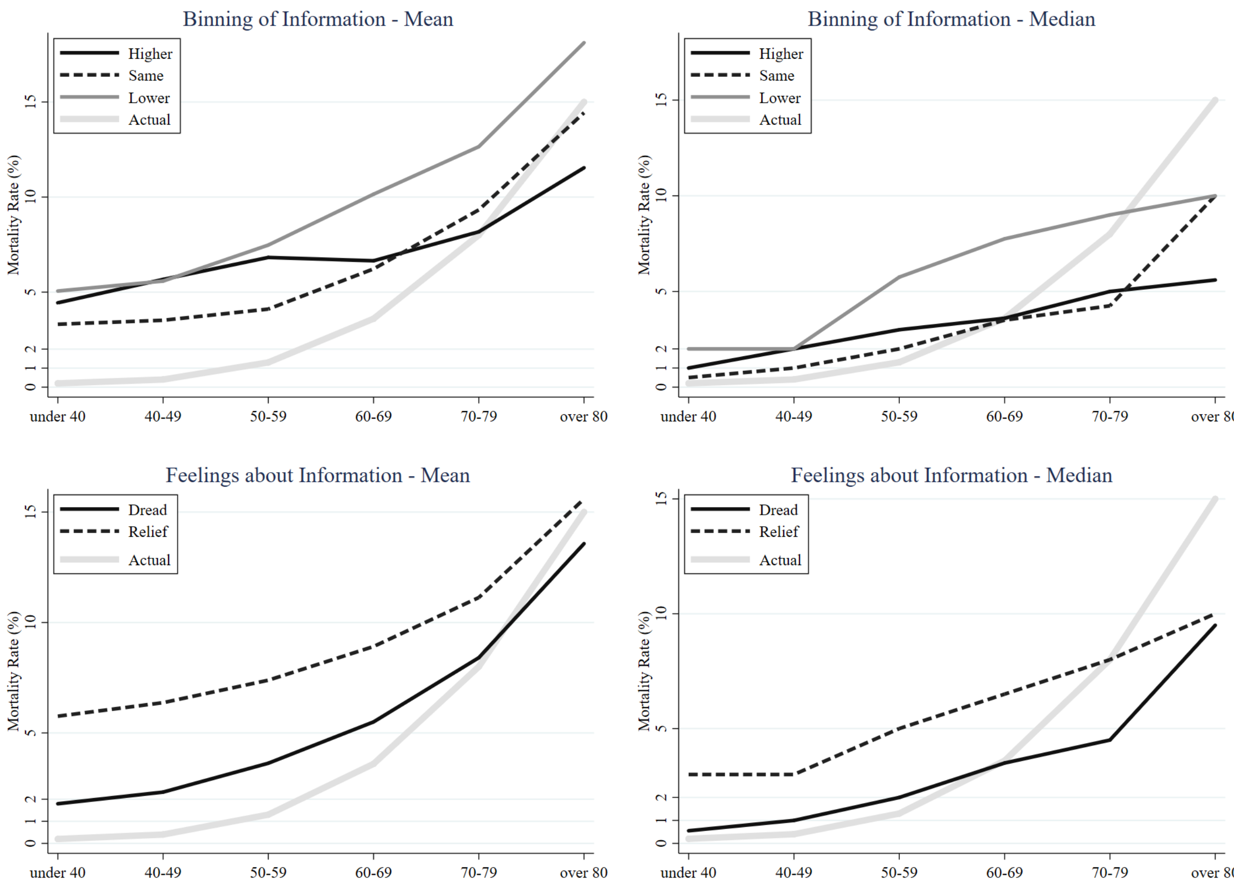
Figure 12 shows initial risk perceptions for respondents categorized by how they later assessed the information received. To draw the age-risk gradients for these people, we note that 43% of respondents indicated that the information was the same as they thought, while 27% said it was higher and 30% said it was lower. We see that those who bin the provided risk information as "lower" than they thought did, in fact, initially have the highest risk perceptions for older people and vice versa for those who chose the "higher" bin. But importantly, we find evidence that a substantial share of respondents report that the information we provided was the "same" as what they thought, even if the true risk was, in many cases, much lower than expected. For instance, the average risk perception for people under 50 is an order of magnitude off, suggesting that many categorize both a 0.2% and 2% mortality risk as "low risk."

In addition, we find that participants' subjective categorizations of risk are in many cases unrelated to their prior beliefs. In the own information treatment, for instance, 72% of those who said that the information was higher than they thought had, in fact, *overestimated* the risk for their own age group initially. This means that their binning of risk updating was objectively incorrect. Likewise, among those who also received information about the risk to older people, 36% in the "higher" bin had already overestimated the risk for older people.

In contrast to the subgroup analysis based on objective risk misperceptions, as determined by us, the researchers (recall Table A8), we find that subjective categorizations of risk belief updating is highly predictive of charitable giving and attention to prevention. Figure 13 shows how prosocial behavior varies within the treatment arm, depending on whether

[13] This initial finding is consistent with a fatalistic response to bad news (as described in (Kerwin, 2018)) or conversely a positive response to good news (Gable and Reis, 2010) and evidence from (Akesson et al., 2020)

Figure 12: Estimates of Mortality Risk Age Gradient by Subjective Perception (Experiment 3)

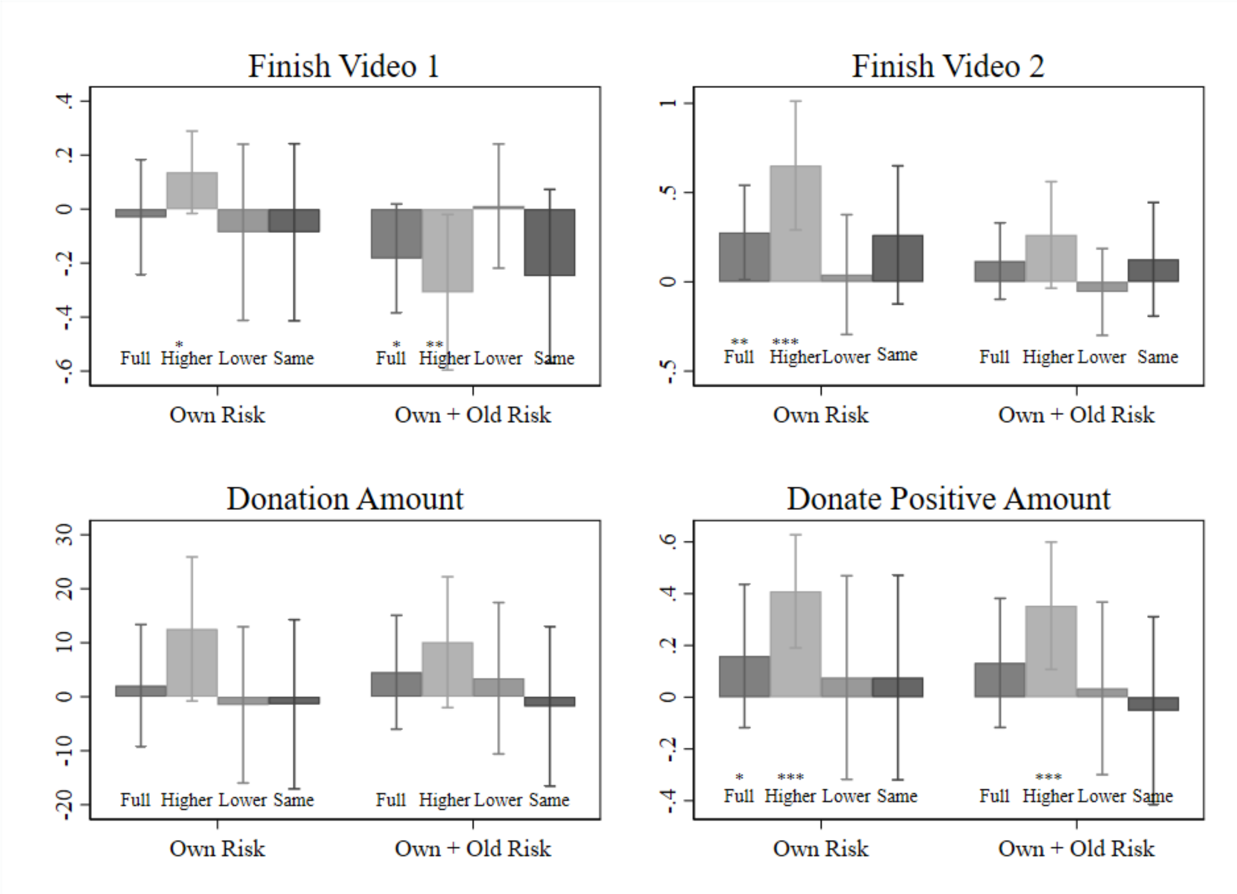


Notes: Estimates of means are based on risk perceptions from Wave 2 wizorized at a 10% level. In the Feelings panel the Dread category includes respondents whose reported feelings at or above the median (of 6) on a 0 (relief) to 10 (dread) scale.

respondents categorized risk information as lower, higher or the same as expected. Interestingly, we find that the group yielding the strongest (i.e., largest and most statistically significant) impacts is the one comprised of those who reported that the risk information was higher than they thought. Further, given that we see little impact among people who said the information was lower or the same, these results suggest that the modest impacts in the original experiment may be due to a muted response among the substantial proportion of respondents whose subjective impression of risk was not shifted even if objectively it appears it should have been (i.e., those people with large "just noticeable differences"). Returning to the underlying social dilemma, it is also important to note that those participants who perceive that they received "good" news do not become less prosocial.

The importance of subjective responses to information is also confirmed by an analysis

Figure 13: Treatment Effects by Perceptions of Information (Experiment 3)



Notes: Each panel reports the coefficients from four separate regressions, one for the full sample and one for each of the three specified binning subgroups. For each subgroup (high, low, or same), the sample is restricted to the control group and the treated respondents who reported that bin and we estimate a regression of the relevant outcome on indicators of for receiving own or own + old risk information. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

of emotional responses. At the end of the manipulation experiment, we asked respondents to report how the information made them feel. The association of these responses with participants *initial* risk perceptions is summarized in the lower panel of Figure 12. The two graphs in this panel show that there are large differences in initial risk perception across respondents who *later* feel relief versus dread upon learning the actual risk and the direction of the difference makes sense. Those who had the highest risk perceptions were more likely to get good news (and feel relief) and those with lower risk perceptions were more likely to learn that the risk was worse than they thought (and feel dread). In fact, the subjective binning of new information and the emotional response to this information are correlated. Those who categorize the risk as higher than they originally thought also feel more dread (average of 7.7 on 0-10 scale) than those who categorize it as the same (average of 5.7) or lower (average of 5.0) as shown in Appendix Figure A3.

As shown in Figure 14, this feeling of dread (as opposed to relief) predicts how much time people invest in watching the videos and their donations to the public good. This suggests that people who were more emotionally impacted by the information were more likely to change their behavior and is consistent with other studies showing that the only robust predictor of positive behavior change (e.g., social distancing, improved hand hygiene) was fear of COVID-19 (Harper et al., 2020).

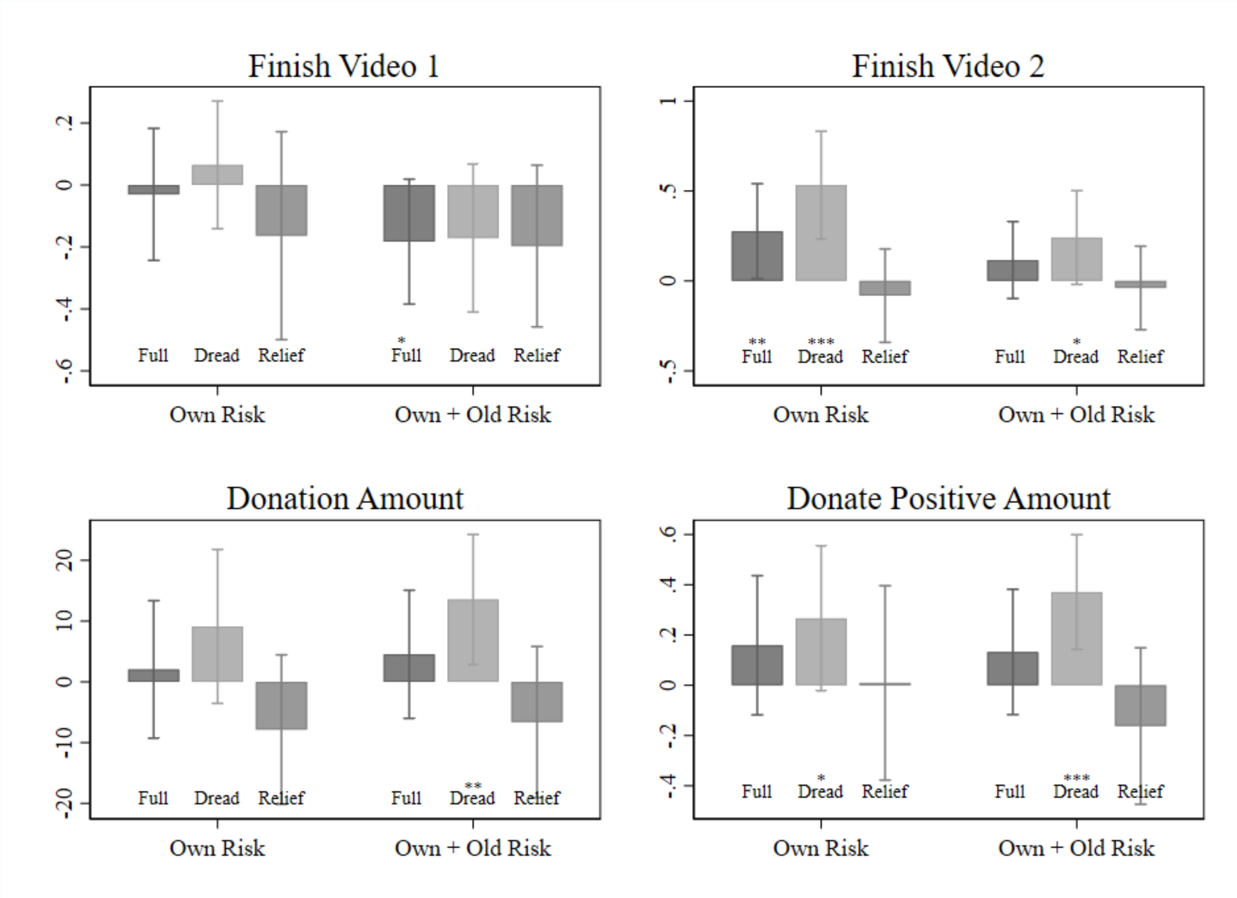
The take away from this section is that the vast majority of respondents reported finding the information treatment credible and relevant, suggesting that lack of relevance and credibility are not driving the muted impacts effects of our debiasing treatment. Based on our third experiment, we conclude that muted treatment effects in the aggregate reflect heterogeneous treatment effects. We find that respondents' subjective reports of how the information treatment changed their broad perception of risk and how the information made them feel are strongly predictive of behavior while our objective measures of how much their perceptions should have changed were not. Respondents who reported that the risk information was much higher than they originally thought and who felt dread upon receiving this information were more likely to donate and watch videos about protecting old people. Those who said the information was the "same" as they originally thought (even if objectively their numerical estimates were very different from the actual information), not surprisingly in retrospect, did not have behavior perceptibly different from the control group.

6 Discussion

Our study highlights some of the difficulties that may hinder the efficacy of officials tasked with presenting important information to the public. Not only is it difficult to capture and maintain the public's attention, when perceptions of the information are subjective and bias is likely, our results highlight the importance of these public officials "thinking slowly" and endeavoring to craft the right message. Put differently, our study indicates that considerable benefits can accrue to matching information, or the presentation of this information, to the anticipated misperceptions or biases likely to arise in the general public. In our case, where the evidence suggests the availability heuristic is likely to affect one's perception of the mortality age-risk gradient, to help prevent the spread of a pandemic, officials should emphasize not just the actual risks to the individual, but also the risks to other, more vulnerable, populations. We find that this can enhance prosocial behavior.

Alongside our main behavioral results, our study suggests important lessons for the design of future experiments. We focus on incentivized measures of prosocial behavior, while collecting self-reports as a touchstone to the existing literature. Given the self-reports are hypothetical and often yield little variation (as others have found, Akesson et al. (2020)), it is not surprising to find our debiasing treatments had little effect, even though sensible effects appear when considering the incentivized outcomes. While it is possible that this is

Figure 14: Treatment Effects by Feelings about Information (Experiment 3)



Notes: Each panel reports the coefficients from three separate regressions, one for the full sample and one for each of the those above and below the median relief-to-dread scale. For each subgroup (relief or dread), the sample is restricted to the control group and the treated respondents who reported those feelings and we estimate a regression of the relevant outcome on indicators of for receiving own or own + old risk information. *** p<0.01, ** p<0.05, * p<0.1

because these outcomes measure different types of prosocial behavior (e.g., mask wearing vs. donations), collecting incentivized measures seems particularly important in studying COVID-19 and other topics with socially desirable behaviors. As important as incentivizing responses, collecting data on subjective assessments of the information treatments was important to explain the mechanisms underlying our main results. This echoes a recent review article by Haaland et al. (2020) who stress the importance of collecting posterior beliefs in information experiments, as it allows researchers to measure whether recipients' pay attention to the information. Our results suggest that in addition to posterior beliefs, researchers should collect data on how people interpret and feel about the information. In our context, the initial misperception of risks does not closely predict recipients' emotional response to and categorization of new information. On top of this, our subjective assessments of information were much more predictive of behavioral changes than the (objective) levels of misestimation.

We conclude by returning to the question that motivated our study: in a setting in which the pandemic is likely to be top of mind and individuals are likely to overestimate their own mortality risk while underestimating the risks to others, can you debias these individuals without creating a social dilemma? In a nutshell, such a social dilemma is not inevitable. Were we to just inform individuals of their own actual risk, we do find the anticipated reduction in preventative measures, however, this effect is attenuated by the careful presentation of information that clarifies the benefits to others of maintaining one's vigilance. In sum, our results confirm that, in the context of COVID-19, providing more detailed information about those who may be most affected by the actions of others, can avoid this social dilemma and achieve the both goals: fostering prosocial behavior and having informed citizens.

References

Abel, M. and W. Brown (2020). Prosocial behavior in the time of COVID-19: The effect of private and public role models.

Abel, M., S. Cole, and B. Zia (2020). Debiasing on a roll: Changing gambling behavior through experiential learning. *The World Bank Economic Review*.

Agha, S. (2003). The impact of a mass media campaign on personal risk perception, perceived self-efficacy and on other behavioural predictors. *AIDS care 15*(6), 749–762.

Akesson, J., S. Ashworth-Hayes, R. Hahn, R. D. Metcalfe, and I. Rasooly (2020). Fatalism, beliefs, and behaviors during the COVID-19 pandemic. Technical report, National Bureau of Economic Research.

Alsan, M. and M. Wanamaker (2018). Tuskegee and the health of black men. *The quarterly journal of economics 133*(1), 407–455.

Bensi, L., R. Nori, E. Strazzari, and F. Giusberti (2003). Vividness in judgements of guilt. *Perceptual and motor skills 97*(3), 1133–1136.

Branas-Garza, P., D. A. Jorrat, A. Alfonso, A. M. Espin, T. García, and J. Kovarik (2020). Exposure to the COVID-19 pandemic and generosity in southern Spain.

Bränström, R., S. Kristjansson, and H. Ullen (2006). Risk perception, optimistic bias, and readiness to change sun related behaviour. *The European Journal of Public Health 16*(5), 492–497.

Brewer, N. T., G. B. Chapman, F. X. Gibbons, M. Gerrard, K. D. McCaul, and N. D. Weinstein (2007). Meta-analysis of the relationship between risk perception and health behavior: the example of vaccination. *Health psychology 26*(2), 136.

Brody, C. J. (1984). Differences by sex in support for nuclear power. *Social forces 63*(1), 209–228.

Broihanne, M.-H., M. Merli, and P. Roger (2014). Overconfidence, risk perception and the risk-taking behavior of finance professionals. *Finance Research Letters 11*(2), 64–73.

Buhrmester, M., T. Kwang, and S. D. Gosling (2016). Amazon's mechanical turk: A new source of inexpensive, yet high-quality data?

Bursztyn, L., A. Rao, C. Roth, and D. Yanagizawa-Drott (2020). Misinformation during a pandemic. *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2020-44).

Butler, G. and A. Mathews (1987). Anticipatory anxiety and risk perception. *Cognitive therapy and research 11*(5).

Campos-Mercade, P., A. Meier, F. Schneider, and E. Wengström (2020). Prosociality predicts health behaviors during the COVID-19 pandemic. *University of Zurich, Department of Economics, Working Paper* (346).

Combs, B. and P. Slovic (1979). Newspaper coverage of causes of death. *Journalism quarterly 56*(4), 837–849.

DeJoy, D. M. (1989). The optimism bias and traffic accident risk perception. *Accident Analysis & Prevention 21*(4), 333–340.

Delavande, A. and H.-P. Kohler (2016). HIV/AIDS-related expectations and risky sexual behaviour in Malawi. *The Review of Economic Studies 83*(1), 118–164.

Dhont, K., A. Van Hiel, and D. De Cremer (2012). Externalities awareness in anticommons dilemmas decreases defective behavior. *Journal of Behavioral Decision Making 25*(3), 228–238.

Dillard, A. J. and J. L. Main (2013). Using a health message with a testimonial to motivate colon cancer screening: Associations with perceived identification and vividness. *Health Education & Behavior 40*(6), 673–682.

Eil, D. and J. M. Rao (2011). The good news-bad news effect: asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics 3*(2), 114–38.

Epley, N., B. Keysar, L. Van Boven, and T. Gilovich (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of personality and social psychology 87*(3), 327.

Erceg, N., M. Ružojčić, and Z. Galic (2020). Misbehaving in the Corona crisis: The role of anxiety and unfounded beliefs.

Falco, P. and S. Zaccagni (2020). Promoting social distancing in a pandemic: Beyond the good intentions.

Fetzer, T., L. Hensel, J. Hermle, and C. Roth (2020). Coronavirus perceptions and economic anxiety. *Review of Economics and Statistics*, 1–36.

Finucane, M. L., P. Slovic, C. K. Mertz, J. Flynn, and T. A. Satterfield (2000). Gender, race, and perceived risk: The white male effect. *Health, risk & society 2*(2), 159–172.

Frederick, S. (2005, December). Cognitive reflection and decision making. *Journal of Economic Perspectives 19*(4), 25–42.

Frh, H. (2017). Risk perception as media effect. *The International Encyclopedia of Media Effects*, 1–8.

Gable, S. L. and H. T. Reis (2010). Good news! capitalizing on positive events in an interpersonal context. In *Advances in experimental social psychology*, Volume 42, pp. 195–257. Elsevier.

Garrett, N., A. M. González-Garzón, L. Foulkes, L. Levita, and T. Sharot (2018). Updating beliefs under perceived threat. *Journal of Neuroscience 38*(36), 7901–7911.

Gong, E. (2015). HIV testing and risky sexual behaviour. *The Economic Journal 125*(582), 32–60.

Gwartney-Gibbs, P. A. and D. H. Lach (1991). Sex differences in attitudes toward nuclear war. *Journal of Peace Research 28*(2), 161–174.

Haaland, I., C. Roth, and J. Wohlfahrt (2020). Designing information provision experiments.

Harper, C. A., L. P. Satchell, D. Fido, and R. D. Latzman (2020). Functional fear predicts public health compliance in the COVID-19 pandemic. *International journal of mental health and addiction.*

Hart, S. G. and L. E. Staveland (1988). Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, Volume 52, pp. 139–183. Elsevier.

Heereman, J. and P. Walla (2011). Stress, uncertainty and decision confidence. *Applied psychophysiology and biofeedback 36*(4), 273–279.

Hesse, B. W., D. E. Nelson, G. L. Kreps, R. T. Croyle, N. K. Arora, B. K. Rimer, and K. Viswanath (2005). Trust and sources of health information: the impact of the internet and its implications for health care providers: findings from the first health information national trends survey. *Archives of internal medicine 165*(22), 2618–2624.

Houghton, S. M., M. Simon, K. Aquino, and C. B. Goldberg (2000). No safety in numbers: Persistence of biases and their effects on team risk perception and team decision making. *Group & Organization Management 25*(4), 325–353.

Johnson, E. D., E. Tubau, and W. De Neys (2014). The unbearable burden of executive load on cognitive reflection: A validation of dual process theory. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, Volume 36.

Kassam, K. S., K. Koslov, and W. B. Mendes (2009). Decisions under distress: Stress profiles influence anchoring and adjustment. *Psychological science 20*(11), 1394–1399.

Keller, C., M. Siegrist, and H. Gutscher (2006). The role of the affect and availability heuristics in risk communication. *Risk analysis 26*(3), 631–639.

Kerwin, J. (2018). Scared straight or scared to death? the effect of risk beliefs on risky behaviors. *The Effect of Risk Beliefs on Risky Behaviors (February 9, 2018).*

Kpanake, L., B. Chauvin, and E. Mullet (2008). Societal risk perception among African villagers without access to the media. *Risk Analysis: An International Journal 28*(1), 193–202.

Levy, M. and J. Tasoff (2016). Exponential-growth bias and lifecycle consumption. *Journal of the European Economic Association 14*(3), 545–583.

Lichtenstein, S., P. Slovic, B. Fischhoff, M. Layman, and B. Combs (1978). Judged frequency of lethal events. *Journal of experimental psychology: Human learning and memory 4*(6), 551.

Lipset, S. M. and W. Schneider (1987). The confidence gap during the Reagan years, 1981-1987. *Political Science Quarterly 102*(1), 1–23.

Matthews, A. K., S. A. Sellergren, C. Manfredi, and M. Williams (2002). Factors influencing medical information seeking among African American cancer patients. *Journal of health communication 7*(3), 205–219.

Merz, Z. C., J. W. Lace, and A. M. Eisenstein (2020). Examining broad intellectual abilities obtained within an mturk internet sample. *Current Psychology*, 1–9.

Morsanyi, K., C. Busdraghi, and C. Primi (2014). Mathematical anxiety is linked to reduced cognitive reflection: a potential road from discomfort in the mathematics classroom to susceptibility to biases. *Behavioral and Brain Functions 10*(1), 1–13.

New York Times (2020). Coronavirus (COVID-19) data in the united states. Technical report. https://github.com/nytimes/covid-19-data.

Nisbett, R. E. and L. Ross (1980). Human inference: Strategies and shortcomings of social judgment.

O'neill, B. (1977). A decision-theory model of danger compensation. *Accident Analysis & Prevention 9*(3), 157–165.

Owen, A. M., K. M. McMillan, A. R. Laird, and E. Bullmore (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human brain mapping 25*(1), 46–59.

Pachur, T., R. Hertwig, and F. Steinmann (2012). How do people judge risks: availability heuristic, affect heuristic, or both? *Journal of Experimental Psychology: Applied 18*(3), 314.

Pennycook, G., J. McPhetres, B. Bago, and D. Rand (2020). Predictors of attitudes and misperceptions about COVID-19 in Canada, the UK, and the US.

Pew Research Center (2014, Oct). Ideological placement of each source's audience. Technical report. https://www.journalism.org/2014/10/21/political-polarization-media-habits/pj_14-10-21_mediapolarization-08/.

Rabin, R. C. (2020, Feb). Why the Coronavirus seems to hit men harder than women. *New York Times*.

Rains, S. A. (2007). Perceptions of traditional information sources and use of the world wide web to seek health information: findings from the health information national trends survey. *Journal of health communication 12*(7), 667–680.

Robb, K. A., A. Miles, J. Campbell, P. Evans, and J. Wardle (2008). Impact of risk information on perceived colorectal cancer risk: A randomized trial. *Journal of Health Psychology 13*(6), 744–753.

Romer, D., K. H. Jamieson, and S. Aday (2003). Television news and the cultivation of fear of crime. *Journal of communication 53*(1), 88–104.

Savage, I. (1993). Demographic influences on risk perceptions. *Risk analysis 13*(4), 413–420.

Schaeffer, M. H. (1989). Environmental stress and individual decision-making: Implications for the patient. *Patient Education and Counseling 13*(3), 221–235.

Schwartz, S. H. (1977). Normative influences on altruism. *Advances in Experimental Social Psychology 10*(1), 221–279.

Shaham, Y., J. E. Singer, and M. H. Schaeffer (1992). Stability/instability of cognitive strategies across tasks determine whether stress will affect judgmental processes'. *Journal of Applied Social Psychology 22*(9), 691–713.

Shedler, J. and M. Manis (1986). Can the availability heuristic explain vividness effects? *Journal of personality and social psychology 51*(1), 26.

Simon, M., S. M. Houghton, and K. Aquino (2000). Cognitive biases, risk perception, and venture formation: How individuals decide to start companies. *Journal of business venturing 15*(2), 113–134.

Simonov, A., S. K. Sacher, J.-P. H. Dubé, and S. Biswas (2020). The persuasive effect of Fox news: non-compliance with social distancing during the COVID-19 pandemic. Technical report, National Bureau of Economic Research.

Sjåstad, H. and J. J. Van Bavel (2020). The best-case heuristic in risk prediction: Hopes and fears in a global health pandemic (COVID-19).

Sjöberg, L. and E. Engelberg (2010). Risk perception and movies: A study of availability as a factor in risk perception. *Risk Analysis: An International Journal 30*(1), 95–106.

Slovic, P. E. (2000). *The perception of risk.* Earthscan publications.

Sobkow, A., J. Traczyk, and T. Zaleskiewicz (2016). The affective bases of risk perception: negative feelings and stress mediate the relationship between mental imagery and risk perception. *Frontiers in psychology 7*, 932.

Stango, V. and J. Zinman (2009). Exponential growth bias and household finance. *The Journal of Finance 64*(6), 2807–2849.

Stanley, M., P. Seli, N. Barr, and K. Peters (2020). Analytic-thinking predicts hoax beliefs and helping behaviors in response to the COVID-19 pandemic.

Tamir, D. I. and J. P. Mitchell (2010). Neural correlates of anchoring-and-adjustment during mentalizing. *Proceedings of the National Academy of Sciences 107*(24), 10827–10832.

Toplak, M. E., R. F. West, and K. E. Stanovich (2011). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & cognition 39*(7), 1275.

Traczyk, J., A. Sobkow, and T. Zaleskiewicz (2015). Affect-laden imagery and risk taking: the mediating role of stress and risk perception. *PloS one 10*(3), e0122226.

Tversky, A. and D. Kahneman (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology 5*(2), 207–232.

Tversky, A. and D. Kahneman (1974). Judgment under uncertainty: Heuristics and biases. *science 185*(4157), 1124–1131.

Verity, R., L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, G. Cuomo-Dannenburg, H. Thompson, P. G. Walker, H. Fu, et al. (2020). Estimates of the severity of Coronavirus disease 2019: a model-based analysis. *The Lancet infectious diseases*.

Wahlberg, A. A. and L. Sjoberg (2000). Risk perception and the media. *Journal of risk research 3*(1), 31–50.

Weinstein, N. D. (1989). Optimistic biases about personal risks. *Science 246*(4935), 1232–1234.

Weinstein, N. D. and W. M. Klein (1995). Resistance of personal risk perceptions to debiasing interventions. *Health psychology 14*(2).

Wilde, G. J. (1982). The theory of risk homeostasis: implications for safety and health. *Risk analysis 2*(4), 209–225.

A Appendix

Table A1: Tests of Rank Ordering of Risk Perceptions by Baseline Characteristics

	Under 40	40-9	50-59	60-69	70-79	Over 80
Respondent Age						
Jonckheere Terpstra Test	0.000	4.91e-10	1.38-08	8.10e-09	6.02e-08	3.64e-07
Test that mean risk perception of:						
under 30 > 30 to 39	0.00523	0.00838	0.0179	0.0899	0.0317	0.00665
40 plus > 30 to 39	5.11e-05	0.00392	0.00400	0.000362	0.000821	0.00591
Test that median risk perception of:						
under 30 > 30 to 39	0.500	0.00610	0.0520	0.00123	0.0846	0.000264
40 plus > 30 to 39	3.76e-06	0.00399	0.000303	0.0163	0.00186	0.0331
Gender						
Test that mean risk perception of:						
Male > female	0.116	0.0517	0.0127	0.0890	0.0615	0.256
Test that median risk perception of:						
Male > female	0.500	0.500	0.0791	0.500	0.0570	0.0576
News Source						
Test that mean risk perception of:						
Liberal > Center and Right	0.0428	0.166	0.252	0.193	0.0894	0.0510
Test that median risk perception of:						
Liberal > Center and Right	4.73e-07	0.00222	0.0287	0.500	0.279	0.251
State Deaths						
Test that mean risk perception of:						
Any Deaths > No Deaths	0.0680	0.00650	0.0224	0.148	0.244	0.174
Test that median risk perception of:						
Any Deaths > No Deaths	0.198	0.500	0.287	0.500	0.279	0.251
Cognitive Score						
Jonckheere Terpstra Test	2.86e-05	7.55e-06	1.17e-05	0.000572	0.0362	0.0833
Test that mean risk perception of:						
Zero correct > One correct	0.0111	0.0190	0.0322	0.0450	0.271	0.178
One > Two	0.105	0.0198	0.0171	0.0282	0.0631	0.0824
Two > Three	0.0186	0.0541	0.0597	0.186	0.375	0.440
Test that median risk perception of:						
Zero correct > One correct	0.500	0.0385	0.0124	0.00334	0.0499	0.0275
One > Two	0.00476	0.0417	0.136	0.253	0.0654	0.0820
Two > Three	0.500	0.500	0.500	0.786	0.726	0.500

Notes: For ranked comparisons of subgroups with three or more categories (example respondent age), p-values of Jonckeree Terpstra tests of rank ordering are reported in the first row. Next, we report the p-value of sequential one-sided tests of each category compared to the next in rank order, first at the means and then at the medians. For comparisons with only two categories (ex. gender) we report the p-value of one-sided tests of the ranking of risk estimates as seen in figures 1 and 2.

Table A2: Baseline Characteristics by Treatment (Experiment 1 and 3)

	N	Sample	Control	Load	Nudge	Manipul.
Age	386	36.2	37.33	35.22	37.02	35.21
Female	386	.37	.35	.36	.43	.33
4 yr college	386	.58	.51	.65*	.56	.61
Liberal	386	.41	.43	.39	.4	.43
Conservative	386	.36	.36	.34	.34	.4
Worried Corona	386	.4	.39	.45	.42	.34
Prevent Corona	386	.36	.37	.41	.36	.31
Spread Virus	386	.81	.84	.73	.83	.84

Notes: Column *Sample* reports average values for the full sample. The remaining columns report average values for the randomly assigned groups. *Worried Corona* reports the share who believe they will contract the virus and *Prevent Corona* reports the share who believe they cannot protect themselves from the virus. Significance is reported for a test of equal means between the control and respective treatment groups. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3: Treatment Effects: Cognitive and Risk Perceptions (Experiment 1)

	Risk Perception					
	Under 40	40-49	50-59	60-69	70-79	Over 80
Cognitive Load	19.037**	22.860**	24.832**	41.089***	57.230***	46.975*
	(9.562)	(9.523)	(9.856)	(12.385)	(15.954)	(24.103)
Observations	185	185	185	185	185	185
R^2	0.021	0.031	0.034	0.057	0.066	0.020
Sample Mean	42.76	49.66	61.54	82.95	109.85	155.45
Std Dev	63.60	63.741	68.23	83.98	105.275	163.08

Notes: Dependent variables measure the mortality risk perceptions (out of 1,000 infected people) for different age groups. All estimations are OLS. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A4: Heterogeneous Treatment Effects by Knowing Victims (Experiment 1)

	Risk Perception						
	Own Risk	Under 40	40-49	50-59	60-69	70-79	over 80
Nudge treat	24.83	14.91	10.62	10.34	1.014	-15.24	-10.79
	(25.52)	(10.17)	(10.42)	(11.51)	(14.51)	(18.27)	(28.47)
Know victim	149.7***	52.47***	42.10**	37.56*	19.02	5.065	33.44
	(41.45)	(14.53)	(14.21)	(15.04)	(16.79)	(19.77)	(33.78)
Nudge x Know Victim	-14.94	-11.79	-0.119	1.413	30.27	30.36	-1.604
	(58.57)	(20.84)	(20.43)	(21.42)	(25.08)	(27.83)	(47.58)
R^2	.139	.117	.098	.075	.052	.0167	.0108
Observations	193.00	193.00	193.00	193.00	193.00	193.00	193.00
Sample Mean	104.54	42.76	49.66	61.54	82.95	109.85	155.45
Standard Deviation	169.158	63.607	63.741	68.229	83.979	105.275	163.082
P-value: joint	0.851	0.864	0.551	0.516	0.128	0.472	0.745

Notes: The table reports risk perceptions for the assignment to the availability nudge treatment, whether people know of a victim, and the interaction of these two variables. Risk perceptions are winsorized at the 10% level. The p-value in the bottom row reports reports on whether the sum of the nudge treatment and the interaction term are different from zero. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A5: Treatment Effects: Attention to Prevention (Experiment 2)

	Video: "*How Protect Yourself*"				Video: "*How Protect Old People*"			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1: Own risk	0.013	0.018	0.013	0.018	-0.089**	-0.077*	-0.089**	-0.077*
	(0.046)	(0.044)	(0.046)	(0.044)	(0.043)	(0.042)	(0.043)	(0.042)
T2+T3: Old risk	0.020	0.025			-0.012	-0.009		
	(0.040)	(0.038)			(0.038)	(0.037)		
T2: Old random risk			0.020	0.019			-0.040	-0.037
			(0.046)	(0.045)			(0.044)	(0.043)
T3: Old close risk			0.021	0.031			0.015	0.018
			(0.046)	(0.045)			(0.045)	(0.044)
Control Variables	N	Y	N	Y	N	Y	N	Y
Observations	928	928	928	928	928	928	928	928
Sample Mean	0.46	0.46	0.46	0.46	0.33	0.33	0.33	0.33
Std Dev	0.50	0.50	0.50	0.50	0.47	0.47	0.47	0.47
T1=T2+T3	0.848	0.865			0.035	0.070		
T1=T2			0.874	0.982			0.247	0.355
T1=T3			0.864	0.787			0.016	0.029
T2=T3			0.990	0.805			0.215	0.210

Notes: The dependent variable in Column (1) through (4) is whether participants finish watching the video describing how people can protect themselves. The dependent variable in Col. (5) through (8) measures whether people finish watching the video on how to protect older people. All estimations are OLS. Control variables include socio-demographic variables reported in Table 1. Robust standard errors are in parentheses. The bottom rows of the table report p-values of test of equal coefficients for the treatment group estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A6: Treatment Effects: Donations (Experiment 2)

	Donation: Cents				Donation: 1=Pos. Amount			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1: Own risk	0.205	0.110	0.205	0.106	0.027	0.024	0.027	0.024
	(1.876)	(1.812)	(1.877)	(1.813)	(0.046)	(0.045)	(0.046)	(0.045)
T2+T3: Old risk	2.111	2.255			0.066*	0.066*		
	(1.646)	(1.590)			(0.040)	(0.039)		
T2: Old random risk			1.835	1.899			0.068	0.066
			(1.908)	(1.832)			(0.046)	(0.045)
T3: Old close risk			2.384	2.603			0.065	0.066
			(1.907)	(1.841)			(0.046)	(0.044)
Control Variables	N	Y	N	Y	N	Y	N	Y
Observations	918	918	918	918	918	918	918	918
Sample Mean	18.71	18.71	18.71	18.71	0.57	0.57	0.57	0.57
Std Dev	20.35	20.35	20.35	20.35	0.49	0.49	0.49	0.49
T1=T2+T3	0.243	0.173			0.325	0.286		
T1=T2			0.390	0.323			0.374	0.357
T1=T3			0.250	0.173			0.416	0.352
T2=T3			0.775	0.702			0.939	0.999

Notes: The dependent variable in Column (1) through (4) measures the amount participants donate (in cents). The dependent variable in Col. (5) through (8) measures whether participants donate a positive amount or not. All estimations are OLS. Control variables include socio-demographic variables reported in Table 1. Robust standard errors are in parentheses. The bottom rows of the table report p-values of test of equal coefficients for the treatment group estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A7: Treatment Effects: Attitudes and Self-Reported Behavior (Experiment 2)

	Cancel Dinner				Wear Mask				Tax for Sick Leave			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
T1: Own risk	-0.030	-0.038	-0.030	-0.038	-0.047	-0.058	-0.047	-0.058	0.051	0.030	0.051	0.030
	(0.034)	(0.034)	(0.034)	(0.034)	(0.044)	(0.043)	(0.044)	(0.043)	(0.044)	(0.043)	(0.044)	(0.043)
T2+T3: Old risk	-0.067**	-0.065**			-0.022	-0.016			0.040	0.032		
	(0.030)	(0.030)			(0.038)	(0.037)			(0.039)	(0.037)		
T2: Old random risk			-0.089**	-0.084**			-0.003	0.002			0.057	0.055
			(0.036)	(0.036)			(0.044)	(0.043)			(0.045)	(0.043)
T3: Old close risk			-0.046	-0.046			-0.040	-0.033			0.024	0.010
			(0.035)	(0.034)			(0.044)	(0.043)			(0.045)	(0.044)
Control Variables	N	Y	N	Y	N	Y	N	Y				
Observations	928	928	928	928	928	928	928	928	928	928	928	928
Sample Mean	0.80	0.80	0.80	0.80	0.64	0.64	0.64	0.64	0.63	0.63	0.63	0.63
Std Dev	0.40	0.40	0.40	0.40	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48
T1=T2+T3												
T1=T2			0.116	0.223			0.319	0.171			0.896	0.558
T1=T3			0.655	0.833			0.883	0.576			0.535	0.648
T2=T3			0.261	0.315			0.398	0.423			0.456	0.302

Notes: The dependent variable in Col 1-4 is a dummy for whether people agree that they would cancel dinner, Col. 5-8 is a dummy for whether people would wear masks (even if they do not have symptoms), Col. 9-12 is a dummy for whether people are willing to pay higher taxes to pay for extended sick leave benefits. Control variables include socio-demographic variables reported in Table 1. Robust standard errors are in parentheses. The bottom rows of the table report p-values of test of equal coefficients for the treatment group estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A8: Heterogeneous Treatment Effects by Risk Misconception (Experiment 2)

	V gen	V old	Donation	Don pos	Dinner	Mask	Tax
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
T1: Own risk	-0.029	-0.120**	-1.947	-0.021	-0.039	-0.070	0.057
	(0.052)	(0.048)	(2.098)	(0.052)	(0.041)	(0.055)	(0.055)
T2+T3: Old risk	0.020	-0.040	0.705	0.032	-0.063*	-0.034	0.023
	(0.045)	(0.044)	(1.853)	(0.046)	(0.036)	(0.047)	(0.048)
Overestimate Own Risk	0.077*	0.015	2.891*	0.067*	-0.112**	0.009	-0.063
	(0.044)	(0.043)	(1.605)	(0.040)	(0.052)	(0.049)	(0.057)
T1 x Overestimate Own	0.093	0.074	4.779**	0.116**	0.157**	0.091	0.118*
	(0.058)	(0.060)	(2.056)	(0.046)	(0.061)	(0.068)	(0.070)
T2+T3 x Overestimate Own	-0.007	0.064	3.218*	0.078*	0.105	0.026	0.090
	(0.053)	(0.053)	(1.888)	(0.044)	(0.064)	(0.063)	(0.068)
Overestimate Old Risk	-0.004	0.002	-1.375	-0.011	0.028	0.027	0.045*
	(0.025)	(0.025)	(0.898)	(0.024)	(0.018)	(0.021)	(0.025)
T1 x Overestimate Old	0.006	-0.024	1.311	0.047	-0.028	-0.031	-0.020
	(0.034)	(0.033)	(1.310)	(0.034)	(0.026)	(0.034)	(0.036)
T2+T3 x Overestimate Old	0.036	0.010	1.166	0.013	-0.047**	-0.020	-0.044
	(0.029)	(0.029)	(1.081)	(0.029)	(0.024)	(0.027)	(0.030)
Observations	927	927	917	917	731	731	731
R-square	0.03	0.02	0.05	0.06	0.02	0.01	0.01
Mean	0.46	0.33	18.71	0.57	0.81	0.63	0.63
Std Deviation	0.499	0.470	20.346	0.495	0.390	0.483	0.483

Notes: This table presents results from the following specification: $y_i = \gamma_0 + \gamma_1 OverOwn_i + \gamma_2 OverOld_i + \gamma_3 OverOwn\,x\,T1_i + \gamma_4 OverOwn\,x\,Told_i + \gamma_5 Over_O ld\,x\,T1_i + \gamma_6 Over_o ld\,x\,Told_i + \gamma_7 T1_i + \gamma_8 Told_i + \epsilon_i$. Variable *OverOwn* and *OverOld* are dummy variables equal to 1 if people overestimate the risk of their own age group and the elderly, respectively. *Told* is a dummy equal to 1 if participants are part of either T2 or T3. The dependent variable in Column (1) and (2) is whether participants watch the general and old person video, respectively. All estimations are OLS. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure A1: Comprehension Check (Experiments 1, 2, and 3)

You said 20 out of 1,000 infected people your age will die from the virus.

This translates into a mortality rate of 2%.

Is this your intended answer?

Yes

No (you can change your reply)

Figure A2: CDC Donation Prompt (Experiments 1, 2, and 3)

CDC Foundation

Together our impact is greater

You have the option to **donate** part of your bonus to the CDC Foundation's **Emergency Response Fund**, which will be used to respond to the public health threat posed by this virus. The funds go to additional support for personal **protective equipment** and critical **response supplies**, which may help to **prevent the spread** of the coronavirus.

How much do you want to donate (if any)?

	0	5	10	15	20	25	30	35	40	45	50
Donation											

Figure A3: Subjective Binning vs. Emotional Response (Experiment 3)