**Harvard Data Science Review • Special Issue 1: COVID-19:
Unprecedented Challenges and Chances**

The Coronavirus Exponential: A Preliminary Investigation Into the Public's Understanding

Alexander Podkul^{1,2} Liberty Vittert^{3,4} Scott Tranter⁵ Alex Alduncin⁵

¹Research Science Team, Morning Consult, District of Columbia, United States of America,

²McCourt School of Public Policy at Georgetown University,

³Harvard Data Science Initiative, Harvard University, Cambridge, Massachusetts, United States of America,

⁴John M. Olin Business School, Washington University in St. Louis, St. Louis, Missouri, United States of America,

⁵Optimus Analytics, District of Columbia, United States of America

Published on: May 14, 2020

DOI: <https://doi.org/10.1162/99608f92.fec69745>

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

ABSTRACT

In this period of global uncertainty, the public is being inundated by information (and misinformation) from news sources, social media, and the community about the spread of COVID-19 and disease more generally. Moreover, unlike with [most widespread news coverage](#), data and models are being used to explain the story by news organizations, health organizations, and governments. The reasons we are presently socially distancing are entrenched in an understanding of exponential growth and flattening the curve. A pair of survey experiments run on Øptimus Analytics' Daily National Tracking Poll explores public statistical literacy by examining their ability to calculate and understand exponential growth. Our results present evidence suggesting that although individuals can face difficulty in calculating exponential growth, they do understand the nature of exponential relationships. These findings may be used to help better ground effective communication strategies aimed at the general public. Future research developing out of these initial survey results will continue to explore public understanding of exponential growth both domestically and abroad.

Keywords: coronavirus, COVID-19, exponential growth, linear growth, public understanding, numerical literacy, statistical literacy

1. Introduction

As the medical community, governments, business, and the public come together to fight the COVID-19 pandemic, data scientists have a special role in helping to communicate what the data actually mean. As part of those efforts, having a clear metric of where the public's understanding of exponential growth lies and the potential to understand which sub-populations need further methods of explanation will be integral to social scientists, governments, and other public service organizations in reaching and helping to incentivize good social distancing practices around COVID-19. The data science community cannot clearly communicate findings or even correctly measure success in doing so until there is a strong measure for where the public's understanding presently stands.

Further, throughout the coronavirus pandemic, the public has been issued numerous guidelines aimed at preventing disease spread. However, for these guidelines to be taken seriously, members of the public need to have an approximate understanding as to the rate of spread (see, e.g., Gigerenzer et al., 2007; Gigerenzer and Edwards, 2003).

Given the onslaught of statistical analyses presented to the public in the media, there is a clear question of how an understanding (or lack thereof) of exponential growth might influence an individual's worry over COVID-19 (or perhaps an individual's worry leading them to learn more about the topic). In one preliminary study conducted during the present pandemic, Sevi et al., found that public opinion support for confinement was unchanged relative to how the data were graphically presented to respondents.

In this study, as an initial investigation into testing statistical literacy¹ among Americans with respect to exponential growth, a pair of survey experiments were tested over an 8-day time period. Broadly, these experiments aim to explore public understanding of both linear and exponential relationships applied to two different contexts—one related to the spread of disease and another using an alternative, lighter subject.

The article proceeds as follows. After describing the survey methodology, we introduce our first survey experiment, which asks respondents to calculate different types of growth regarding disease spread in an open-ended question format. Next, we present the second survey experiment, which exposes respondents to a data visualization outlining various trends and asks them to choose which line best describes a growth situation presented to them. Both questions in the survey are conducted by split-testing where half of the survey respondents are randomly assigned to linear growth treatment and the other half are assigned to the exponential growth treatment. By using a split-test design for these two survey experiments, we can best understand respondents' *initial* assessment to trends being described to them. Following a discussion of the survey experiment findings, this paper tests whether understanding exponential growth is associated with increased/decreased worry about the spread of coronavirus. Finally, we briefly explore future research avenues as follow-ups to this analysis.

2. Polling Experiments on Linear vs. Exponential Growth

2.1. Survey Design

The survey used in this analysis was fielded from March 27, 2020 to April 3, 2020 using an online web panel hosted by Dynata and coordinated by Øptimus Analytics.² Throughout the course of the 8 days in the field, 2,312 respondents were interviewed regarding a variety of questions related to COVID-19 and political topics. Recruited panel respondents were contacted according to joint target quotas, which were calculated using the most recent U.S. Census Bureau's Current Population Survey (CPS) by age group, Census region, and gender. Final survey results were weighted to be representative of U.S. adults in two steps: first, using post-stratification to account for daily joint quota non-response and, second, by an iterative proportional fitting algorithm to weight the final sample according to CPS marginal distributions on gender (Male/Female), race (White/Black/Hispanic/Other), education (Less than HS Diploma/Some College or Associate's Degree/College Degree/Post-Graduate Degree), age group (18–34/35–44/45–64/65+), and Census region (Northeast/Midwest/South/West) (Pasek, 2010). Although the present project simply seeks to report the findings for the survey's respondents, these figures are weighted to be representative the all U.S. adults, given the general errors and constraints associated with any type of sampling. While the margin of error for the overall topline of the survey is ± 2.2 percentage points (at 95% confidence) with a design effect of 1.19 (Kish, 1965), subpopulation margins of error are recalculated as necessary.³

We emphasize that our web-panel-based survey is not a probabilistic sample. The use of nonprobability sampling for reporting population figures is common among both industry and academic polling. For example,

survey firms such as YouGov and Ipsos regularly report findings derived from nonprobability panels. Many academic polls, such as the recent 50-State COVID-19 Survey out of Northeastern University, Harvard Kennedy School, and Rutgers University (Lazer, Baum, Ognyanova, & Volpe, 2020) or the ongoing Democracy Fund + UCLA Nationscape study (Tausanovitch, Vavreck, Reny, Hayes, & Rudkin, 2019) also rely exclusively on nonprobability panels. As an effect of the methodology described above—viz. the use of quota targeting on the Dynata panel and weighting adjustments—the survey presented in this paper reflects common contemporary polling practices. We understand well the inherent discussion concerning non-probability sampling and have done our best to mitigate and assuage concerns. We are therefore comfortable using these individual experimental results as the best snapshot available for the public's understanding of these issues at this time. We, ourselves, are continuing a further exploration/continued polling and encourage others to replicate this experimental design for further results.

2.2. Calculating Growth

The first survey experiment asks respondents a word problem-style survey question to assess their ability in calculating different types of growth. In this experiment, respondents were randomly assigned to answer one of the two following open-ended survey questions:

1. *Linear Growth*: If 10 people were newly diagnosed with a disease each day, approximately how many people would be diagnosed with the disease at the end of 10 days?
2. *Exponential Growth*: If 10 people were newly diagnosed with a disease each day and each infected person spread the disease to 10 new patients the following day, approximately how many people would be diagnosed with the disease at the end of 10 days if this trend continued?

Figure 1. Histogram displaying the responses to Question 1 for each question version where bar heights represent the count of responses within each bin. The vertical dashed line on each version represents the ‘correct’ response to each question. Seven outlying responses greater than 10^{15} were excluded for visual interpretation.

The paneled histograms in Figure 1 display descriptive findings across the two treatment groups. The Linear Growth test finds that the majority (73%) of respondents were able to correctly identify that 100 people would be newly diagnosed with a disease at the end of 10 days. As the top histogram notes, there is a significant concentration of responses near 10^2 with relatively few respondents vastly over-estimating the spread of disease given the stipulated linear trend. The bottom panel in Figure 1 identifies a much wider variance in the responses to the Exponential Growth treatment version. Overall, the exponential test finds far fewer respondents were able to correctly identify the trend. Within our sample, 9 out of 10 of respondents selected a value lower than 10^{10} , which indicates an *underestimate* of how quickly exponential growth can spread considering the parameters noted in the question text.

While epidemiologists—or even the casual newsreader—would state that the question, as presented, oversimplifies the nature of disease spread, this precise question wording was chosen to explore the public’s

ability in calculating linear and exponential growth. The question text for this project was carefully selected to isolate the focus of the question on the trend described and ridding it of distractions such as recontact and other epidemiological factors. Additionally, the project purposely selected the presented parameters for two reasons. First, by providing round numbers (10 days, 10 people), the calculative burden placed on respondents was relatively minor. Though more realistic figures could have been used (such as choosing a rate that doubled every two days), that would have required much more work from the respondents taking this web survey. Second, although the parameters are unrealistic, they are again meant to demonstrate how quickly a trend can grow given fairly basic numbers. We invite future work specifically interested in the public's understanding of COVID-19 spread to adapt the framework used in this survey experiment with the proper parameters.

This first survey experiment finds a general ability to calculate linear trends fairly well. However, when it comes to calculating exponential trends, respondents fare much worse.

2.3. Understanding Growth

Realizing that calculating linear and exponential growth may be a steep calculative burden placed on survey respondents, this project also sought out to test whether survey respondents can properly characterize growth trends, even if they cannot directly calculate them. To assess *understanding*, a second split-sample survey experiment was conducted. Following the framework of the first question, this second survey experiment also randomly assigned respondents to answer a question about linear growth or exponential growth. However, in this second experiment there were two main differences: first, respondents were asked to characterize a trend rather than directly calculate the outcomes from a trend; and, second, respondents were shown a visual aid (presented in Figure 2), which displayed three different options for the trend described in the question text. Additionally, this question did not ask respondents about the spread of disease⁴ but rather the spread of a viral joke. While presented Figure 2, respondents were also shown one of the following two multiple choice questions:

1. *Linear Growth*: If a person text messaged a joke to 2 new friends each day for a week, which line on the above graph would best describe the total number of people who received the joke after 7 days?
2. *Exponential Growth*: If a person text messaged a joke to 2 friends and a day later both of those friends texted the joke to two new friends and this trend continued for a week, which line on the above graph would best describe the total number of people who received the joke after 7 days?

Figure 2. Visual aid given to Question 2 respondents for both the Linear Growth and Exponential Growth versions.

In Figure 2, line A represents the linear trend described in the Linear Growth version and line C represents the exponential trend described in the Exponential Growth version. Line B, though not a correct answer in either of the question versions, describes an exponential trend; however, it shows a trend significantly slower than the one described in the Exponential Growth section.

Figure 3. Responses to the second survey experiment across both versions. Maroon-lled bars (choice A in the Linear Growth version and choice C in the Exponential Growth version) represent the correct answer to each question. Bar ranges represent question margins of error according to 95% con dence and have been adjusted to account for design eects due to weighting.

Table 1. Overall question responses by experimental version.

	Unweighted (%)			Weighted (%)		
	A	B	C	A	B	C
Linear Growth	46.0	27.3	26.7	46.4	27.7	26.0
Exponential Growth	10.3	31.2	58.5	11.2	31.5	57.3

The findings from Question 2 are displayed in and identify that a plurality of respondents in both treatments selected the correct answer to each Version. displays the weighted and unweighted percentages answering each question and finds that 46% of respondents in the linear treatment group were able to correctly identify the linear trend on the chart (margin of error: $\pm 3.1\%$) and 57% of respondents in the exponential treatment were able to correctly identify the exponential trend, with another 32% choosing the incorrect linear trend ($\pm 3.2\%$).

Perhaps unsurprisingly, this survey experiment finds that when presented with a data visualization, respondents are much more likely to properly estimate exponential growth. However, it was unexpected that respondents

were *more likely* to correctly answer the exponential growth treatment compared to the linear growth treatment. While there may be a number of potential reasons—such as the linear trend appearing too flat in the graphic or the current inundation of exponential graphics appearing in the popular press amidst the COVID-19 pandemic.

Taken together, the findings of the two survey experiments suggest that while individuals are unable or unwilling to accurately calculate exponential growth, a majority of them are able to understand the trend when paired with a graphic.

3. Numeracy's (Null) Effects on Disease Worry

In order to connect the survey experiment to broader discussions surrounding COVID-19, an additional analysis sought to test whether understanding exponential growth was connected to respondents' worry about COVID-19. Explicitly, this analysis sought to compare how respondents performed on the second survey experiment with how worried they were about contracting COVID-19. The motivation for this analysis hypothesizes that those who better understood exponential growth (especially those respondents exposed to the exponential growth question) might be more worried considering they properly understood how quickly this trend can grow. Prior research shows that correct risk assessments are informed by properly perceiving the attributes of the hazard (Slovic, Fischhoff, & Lichtenstein, 1980).

This analysis tests whether exponential understanding is associated with a pre-treatment question asked earlier in the survey, which asked respondents How worried are you about the possibility that you or your family members may contract the COVID-19? In this analysis, respondents were coded as Worried if they selected either Very worried or Somewhat worried on the survey and coded as Not worried if they selected either Not too worried or Not worried at all.⁵

To analyze the relationship, a probit model estimated respondent propensity to worry about the coronavirus pandemic conditional on their responses to the survey experiment presented above. In the first model, only the treatment and response terms were included but, in the fuller specification, dummy variable controls were included for the values within each race, gender, education age group, and U.S. Census region. The full model specification and coefficient estimates are displayed in the Appendix, which also reports balance tables presenting the treatment breakdown by demographic subgroup.

Figure 4. Predicted probabilities for $\Pr(\text{Worry} = 1)$ estimated from the full-specification probit model presented in Table 2 across Question 2 experimental versions and broken out by question responses. In calculating predicted probabilities, demographic variables were held to their modes. Vertical point ranges represent two standard errors above and below each probability estimate.

For ease of interpretation, predicted probabilities are plotted across treatment groups and responses in Figure 4 for the fuller model specification while holding the demographic controls at their modes. Properly understanding exponential relationships does not appear to be associated with an increased worry about the COVID-19 pandemic. While further research can develop the relationship between pandemic worry and exponential understanding, these preliminary correlative findings fail to demonstrate any clear connection between being able to understand exponential patterns with being worried about disease outbreak.

This null result merits discussion. Cognitive biases—in the way that individuals misunderstand exponential growth—should, in theory, affect public behavior and their risk assessments (see Slovic, Fischhoff, &

Lichtenstein, 1980). However, these data would suggest otherwise—or that the effect is sufficiently small to not be detected in a study this size. It is possible (and likely) that individuals' worried about the spread of COVID-19 are being driven by factors beyond understanding how fast the disease can spread—its perceived mortality or severity, local orders, elite cues, amount of media consumption, and even political biases can all be other variables that feed into an individual's risk assessment. We have explored one specific attribute that would lead an individual to grow concerned about the spread of COVID-19, but do not speak to these other factors in this study.

3.1. Exploring Consistency

Figure 5. Comparison of responses to Question 1 and Question 2. Darker filled points represent respondents that were 'correct' on both Question 1 and Question 2.

Finally, consistency in numerical understanding was explored. Figure 5 displays the responses to the two survey experiments faceted by the version of the question that was presented to each respondent. Unsurprisingly, those respondents who were given both linear treatments, performed the best with 45% of that group getting the correct answer to both question. Those who were given the linear treatment of the open-

ended question and the exponential treatment for the multiple choice question performed similarly well, at about 42.5%. However, both groups that were exposed to the exponential treatment in the first survey experiment performed significantly worse ($\sim 5\%$ in each group).

4. Conclusion

With confusing medical terms, a lack of real data, and misinformation flying across the internet and into our living rooms, it can be very difficult for the general population to understand the scale and ramifications of the current global pandemic as it unfolds. Statistical models detail potential trajectories for disease spread. These models help to inform decision-makers on the scale and costs associated with fighting the spread of the virus and present an overall picture of what is at stake. This has led to the current estimates of 5- or 6-digit deaths in the United States, which puts COVID-19 on track to be one of the leading causes of death in the United States in 2020.

Beyond decision-makers, these models aim to help individuals understand the risk to themselves and their families. Inherent in these models is a sense of how exponential growth operates— that while there may have only been 100 cases this week, that does not rule out the possibility of thousands of cases next week. However, our initial survey presents evidence that suggests a majority of respondents are unable to fully grasp the scale of exponential growth patterns on their own— but a slight majority are able to when given a visual aid and only three potential options. This implies that while the information produced by models is disseminated broadly, only a small number of individuals are able to fully understand and internalize the full scale. Most respondents have a general understanding of the potential— but aided in connecting the dots between the curves they see on television or other media and the specific circumstances. Our research suggests that people understand exponential growth at a basic level but do not know how to calculate it in scale— which may help to explain why social distancing was initially slow and lackadaisical. Increased public awareness programs on spread of such diseases would help pandemic mitigation efforts to reduce the rate of infections and, ultimately, deaths.

It is clear that the topics explored in this initial research piece are rather nuanced. It is also clear that further work exploring the public's understanding of statistical literacy on exponential growth generally and the spread of COVID-19 more specifically is necessary. The null results on the effects of exponential understanding on COVID-19 worries in particular could use corroboration and a study with higher sample size would be better able to parse out smaller effects. Sub-population analysis is also limited with the current data-set, and would be useful in helping to identify which groups could most benefit from additional public awareness programs. In an effort to spur further research, we ourselves are taking two immediate next steps.

First, we matched our initial survey to a current national consumer file and modeled the entire U.S. adult population on their likelihood to not understand exponential growth patterns. This produces a look-alike universe of the respondents who were most likely to benefit from additional public awareness on the exponential spread of disease. This is an individual-level model covering most American adults, and work is

ongoing to refine and specify the model.⁶ Second, in order to collect more data for this model and to do further analysis on sub-populations, we are fielding additional survey experiments using the same sampling methodology and surveying process as the outbreak continues.

This period of time, where Americans are being fed information on exponential growth of disease spread, presents a unique set of circumstances researchers can use to better understand how and why individuals learn about such concepts. One often laments that the general population does not have the time or resources to understand the data science behind social scientific concepts—this is a rare scenario in which most now have both.

Acknowledgments

We would like to thank a host of individuals who responded with great speed, urgency, and thoughtfulness with ideas on the survey questions including Stephen Ansolabehere, David Hand, David Spiegelhalter and his team (including Gabriel Recchia, Sarah Dryhurst, Claudia Schneider), Mine Cetinkaya-Rundel, Matthew Marolda, Eugene Soltes, Dustin Tingley, Radu Craiu, and Nathan Sanders.

Disclosure Statement

Three of this article's authors are currently employed by Øptimus Analytics, a data science firm specializing in predictive modeling that provided the survey data analyzed in this project.

References

- Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International statistical review*, 70(1), 1–25. <https://doi.org/10.1111/j.1751-5823.2002.tb00336.x>
- Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2007). Helping doctors and patients make sense of health statistics. *Psychological science in the public interest*, 8(2), 53–96. <https://doi.org/10.1111/j.1539-6053.2008.00033.x>
- Gigerenzer, G., & Edwards, A. (2003). Simple tools for understanding risks: from innumeracy to insight. *BMJ*, 327(7417), 741–744. <https://doi.org/10.1136/bmj.327.7417.741>
- Kish, L. (1965). *Survey Sampling*. J. Wiley.
- Lazer, D., Baum, M.A., Ognyanova, K., & Della Volpe, J. (2020). The state of the nation: A 50-state Covid-19 survey. *The COVID-19 Consortium for Understanding the Public's Policy Preferences Across States*. <https://www.politico.com/f/?id=00000171-cda6-d7a8-ad7b-ffefc6230000>

Pasek, J. (2010). ANES Weighting Algorithm: A Description. <https://web.stanford.edu/group/iriss/cgi-bin/anesrake/resources/RakingDescription.pdf>

Sevi, S., Mendoza Avina, M., Peloquin-Skulski, G., Heisbourg, E., Vegas, P., Coulombe, M., Arel-Bundock, V., Loewen, P., & Blais, A. (2020). Logarithmic vs. linear visualizations of COVID-19 cases do not affect citizens' support for confinement. *Canadian Journal of Political Science*, 53(2), 385–390. <https://doi.org/10.1017/S000842392000030X>

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1980). *Facts and fears: Understanding perceived risk*. In R. C. Schwing, & W. A. Albers (Eds.), *Societal Risk Assessment* (pp. 181–216). Springer. https://doi.org/10.1007/978-1-4899-0445-4_9

Tausanovitch, C., Vavreck, L., Reny, T., Rossell Hayes, A., & Rudkin, A. (2019). Democracy fund + UCLA nationscape methodology and representativeness assessment. <https://www.voterstudygroup.org/uploads/reports/Data/NS-Methodology-Representativeness-Assessment.pdf>

Appendix

Table A1: Demographics by treatment group – Question 1.

	Weighted (%)		Unweighted (%)	
	Linear	Exponential	Linear	Exponential
Race				
White	50.2	49.8	50.6	49.4
Black or African American	47.6	52.4	49.5	50.5
Hispanic	53.0	47.0	53.8	46.2
Other	47.5	52.5	47.8	52.2
Education				
HS or Less	49.7	50.3	50.7	49.3

Some College or Associate	50.4	49.6	50.3	49.7
College Grad	52.4	47.6	52.0	48.0
Post-Grad	47.0	53.0	47.2	52.8
Gender				
Male	49.2	50.8	50.2	49.8
Female	51.0	49.0	50.7	49.3
Region				
Northeast	47.3	52.7	47.8	52.2
Midwest	54.1	45.9	54.5	45.5
South	47.6	52.4	48.0	52.0
West	52.9	47.1	53.0	47.0
Age Group				
18-34	50.1	49.9	51.6	48.4
35-44	54.5	45.5	53.3	46.7
45-64	49.5	50.5	49.5	50.5
65+	47.6	52.4	48.5	51.5

Table A2. Demographics by treatment group - Question 2.

	Weighted (%)		Unweighted (%)	
	Linear	Exponential	Linear	Exponential
Race				
White	51.7	48.3	51.3	48.7

Black or African American	49.0	51.0	48.0	52.0
Hispanic	49.0	51.0	48.5	51.5
Other	49.3	50.7	50.3	49.7
Education				
HS or Less	52.1	47.9	51.1	48.9
Some College or Associate	49.8	50.2	50.5	49.5
College Grad	50.2	49.8	49.3	50.7
Post-Grad	50.0	50.0	51.1	48.9
Gender				
Male	51.7	48.3	51.6	48.4
Female	49.9	50.1	49.7	50.3
Region				
Northeast	46.9	53.1	46.6	53.4
Midwest	50.7	49.3	51.6	48.4
South	52.4	47.6	51.7	48.3
West	51.3	48.7	51.1	48.9
Age Group				
18-34	50.2	49.8	49.3	50.7
35-44	52.5	47.5	51.2	48.8
45-64	48.2	51.8	48.4	51.6
65+	54.8	45.2	55.2	44.8

Table A3. Probit model estimating whether understanding growth effects (as measured by the second survey experiment) is related to a respondent's worry about the COVID-19 outbreak.

	Coronavirus Worry = 1	
Intercept	0.75***	1.00***
	(0.08)	(0.17)
Survey Experiment		
Version B	-0.14	-.020
	(0.18)	(0.19)
Response(B)	0.14	0.11
	(0.14)	(0.15)
Response (C)	0.05	0.02
	(0.14)	(0.14)
Version B * Response (B)	0.04	0.07
	(0.24)	(0.25)
Version B * Response (C)	0.01	0.04
	(0.23)	(0.23)
Race		
Black or African American		0.03
		(0.14)
Hispanic		0.21
		(0.12)
Other		0.07

		(0.16)
Gender		
Female		0.10
		(0.09)
Education		
Some College or Associate		0.01
		(0.10)
College Grad		-0.13
		(0.12)
Post-Grad		0.22
		(0.16)
Age Group		
35-44		-0.14
		(0.13)
45-64		-0.07
		(0.11)
65+		-0.22
		(0.12)
Region		
Midwest		-0.17
		(0.14)
South		-0.29*

		(0.13)
West		-0.29*
		(0.14)
Log Likelihood	-673.36	-651.42
Num. Obs.	1121	1103
*** p < 0.001, ** p < 0.01, * p < 0.05		

©2020 Alexander Podkul, Liberty Vittert, Scott Tranter, and Alex Alduncin. This article is licensed under a Creative Commons Attribution (CC BY 4.0) [International license](https://creativecommons.org/licenses/by/4.0/), except where otherwise indicated with respect to particular material included in the article.

Footnotes

1. Specifically, this study investigates a key element of the “statistical knowledge base” necessary for statistical literacy as outlined in Gal (2002). [↵](#)
2. Øptimus Analytics is a member of the American Association for Public Opinion Research (AAPOR) Transparency Initiative. Additional information regarding Optimus surveys can be found at: <https://github.com/optimus-forecasting-and-polling>. [↵](#)
3. Additional materials regarding the survey experiment can be found at: <https://github.com/optimus-forecasting-and-polling/HDSR-Paper-Materials>. [↵](#)
4. Although the first survey experiment does not directly mention COVID-19 so as to not confuse respondents, we nevertheless acknowledge that COVID-19 is likely a top-of-the-head consideration. [↵](#)
5. Respondents who responded “Someone in my family is already infected” (< 1%) or refused to respond to the question (< 1%) were dropped from this analysis. [↵](#)
6. More information on the Øptimus model can be found on the [Github page](#) for this project. [↵](#)