Journal of Experimental Psychology: Applied


ISSN: 1076-898X

2021, Vol. 27, No. 4, 632–656
https://doi.org/10.1037/xap0000403

# Math Matters: A Novel, Brief Educational Intervention Decreases Whole Number Bias When Reasoning About COVID-19

Clarissa A. Thompson[1], Jennifer M. Taber[1], Pooja G. Sidney[2], Charles J. Fitzsimmons[1], Marta K. Mielicki[1], Percival G. Matthews[3], Erika A. Schemmel[1], Nicolle Simonovic[1], Jeremy L. Foust[1], Pallavi Aurora[1], David J. Disabato[1], T. H. Stanley Seah[1], Lauren K. Schiller[4], and Karin G. Coifman[1]
[1] Department of Psychological Sciences, Kent State University
[2] Department of Psychology, University of Kentucky
[3] Department of Educational Psychology, University of Wisconsin-Madison
[4] Teachers College, Columbia University

At the onset of the coronavirus disease (COVID-19) global pandemic, our interdisciplinary team hypothesized that a mathematical misconception—whole number bias (WNB)—contributed to beliefs that COVID-19 was less fatal than the flu. We created a brief online educational intervention for adults, leveraging evidence-based cognitive science research, to promote accurate understanding of rational numbers related to COVID-19. Participants from a Qualtrics panel ($N = 1{,}297$; 75% White) were randomly assigned to an intervention or control condition, solved health-related math problems, and subsequently completed 10 days of daily diaries in which health cognitions and affect were assessed. Participants who engaged with the intervention, relative to those in the control condition, were more accurate and less likely to explicitly mention WNB errors in their strategy reports as they solved COVID-19-related math problems. Math anxiety was positively associated with risk perceptions, worry, and negative affect immediately after the intervention and across the daily diaries. These results extend the benefits of worked examples in a practically relevant domain. Ameliorating WNB errors could not only help people think more accurately about COVID-19 statistics expressed as rational numbers, but also about novel future health crises, or any other context that involves information expressed as rational numbers.

***Public Significance Statement***

In late March 2020 at the beginning of the coronavirus disease (COVID-19) pandemic, we conducted a randomized controlled trial in which a large panel of U.S. adults ($N = 1{,}297$) was randomly assigned to either an educational intervention or a control condition. Those in the educational intervention learned how to accurately compare case-fatality rates for the flu versus COVID-19 by engaging with a brief, online tutorial, which taught them step-by-step how to divide the number of deaths by the number of cases and then compare to find the most fatal virus. The training decreased the likelihood that people mistakenly focused just on the number of deaths, which would have led them to the mistaken conclusion that the flu was more fatal than COVID-19.

*Keywords:* whole number bias, COVID-19, learning, magnitude understanding, worked example

*Supplemental materials:* https://doi.org/10.1037/xap0000403.supp



In early 2020, COVID-19 became a global pandemic, and the World Health Organization (WHO, 2020) warned of an "infodemic." Information, misinformation, and disinformation about the severity of coronavirus disease (COVID-19) have constantly evolved throughout the course of the pandemic. The lack of clear and consistent guidelines about how to curb the spread of the virus at the local, state,

Clarissa A. Thompson https://orcid.org/0000-0001-8758-3218
Jennifer M. Taber https://orcid.org/0000-0003-3285-4871
Pooja G. Sidney https://orcid.org/0000-0003-3719-8543
Charles J. Fitzsimmons https://orcid.org/0000-0002-2529-2164
Marta K. Mielicki https://orcid.org/0000-0001-9522-2180
Percival G. Matthews https://orcid.org/0000-0001-6265-813X
Nicolle Simonovic https://orcid.org/0000-0003-0384-7103
Jeremy L. Foust https://orcid.org/0000-0002-5386-7746
Pallavi Aurora https://orcid.org/0000-0002-3038-8400
David J. Disabato https://orcid.org/0000-0001-7094-4996
T. H. Stanley Seah https://orcid.org/0000-0002-8042-8319
Lauren K. Schiller https://orcid.org/0000-0002-9986-0224
Karin G. Coifman https://orcid.org/0000-0002-2372-0081

Marta K. Mielicki is now at the Rutgers Center for Cognitive Science, Piscataway, New Jersey, United States.

Erika A. Schemmel is now at Department of Psychology, Portland State University.

This research was supported in part by the U.S. Department of Education Institute of Education Sciences Grants R305A160295 and R305U200004 to C. A. Thompson at Kent State University. This study was preregistered on OSF (https://osf.io/9hc7d). All data files and analytic scripts are available on OSF (https://osf.io/fthm3/files/).

Correspondence concerning this article should be addressed to Clarissa A. Thompson, Department of Psychological Sciences, Kent State University, 600 Hilltop Drive, Kent, OH 44240, United States. Email: cthomp77@kent.edu

national, and global levels likely reduced engagement in recommended health behaviors (e.g., wearing masks; Mills et al., 2020).

COVID-19 statistics are commonly presented as rational numbers—whole number frequencies, fractions, or percentages—that are easily misunderstood. In the current experiment, we first demonstrated that misunderstanding of COVID fatality rates is due, in part, to whole number bias (WNB), a common mathematical misconception in which participants focus on whole number numerator and denominator components and not the holistic magnitude of the fractions (Ni & Zhou, 2005). We tested whether an online educational intervention aimed at reducing WNB improved adults' mathematical reasoning about health statistics comparing COVID-19 and the flu. Then, we explored the downstream effects of the intervention on perceived risk, worry, affect, and health behaviors across 10 days of daily diaries. One reason to focus on rational number understanding as a potential predictor of health cognitions about COVID-19 is that rational number understanding is malleable (e.g., Fazio et al., 2016; Fuchs et al., 2016; Sidney, Thompson, et al., 2019; Yu et al., 2020).

WNB occurs when people incorrectly use whole number knowledge to reason about rational numbers (Alibali & Sidney, 2015; Fitzsimmons et al., 2020; Ni & Zhou, 2005; Siegler et al., 2011; Thompson & Opfer, 2008). For example, for the ratio 15/30, people may think about the numerator, 15, or the denominator, 30, as whole numbers in isolation rather than as parts of a ratio (15/30 = ½ = 50% = 0.50). People of different ages (Alibali & Sidney, 2015; Braithwaite & Siegler, 2018; Fazio et al., 2016; Fitzsimmons et al., 2020; Ni & Zhou, 2005; Opfer & Devries, 2008), expertise levels (Obersteiner et al., 2013), and cultures (Alonso-Diaz et al., 2019; DeWolf & Vosniadou, 2015; Gómez et al., 2015; Van Hoof et al., 2020) make WNB errors. Reasoning about the magnitude of ratios compared to whole numbers is more effortful, error-prone, and time-consuming (e.g., Fazio et al., 2014; Siegler et al., 2011; Yu et al., 2020). According to dual-processing theories, people must first inhibit automatic, over-practiced whole-number processes before engaging in effortful, strategic rational number processes (Vamvakoussi, 2015; Vamvakoussi & Vosniadou, 2010; Vosniadou, 2014). Thus, WNB can occur when whole-number processing is not inhibited.

The impetus for this study was our repeated observation of a specific way of communicating COVID-19 statistics in the media in early 2020 (e.g., Faust, 2020; Faust & Del Rio, 2020a, 2020b; Rettner, 2020; Walker, 2020; Yan, 2020) that we hypothesized and would lead people to commit WNB errors. Even though media sources showed the number of deaths and the number of cases together (as in the left two columns of Table 1), they frequently compared only the total number of COVID-19 versus flu deaths. For example, one popular press article was titled, *The flu has killed far more people than coronavirus. So why all the frenzy about COVID-19?* (Netburn, 2020). The article went on to indicate, "The flu has killed tens of thousands more people ... So why is everyone freaking out about the coronavirus?" It is true that the vast majority of people live through flu season every year, with only a small proportion of people dying. However, considering only the absolute number of deaths (22,000 flu deaths > 9,318 COVID-19 deaths in mid-March; Table 1) or the absolute number of people infected with either virus (36,000,000 flu cases > 227,743 COVID-19 cases; Table 1) can lead to the mistaken conclusion that people who get COVID-19 are *less* likely to die than people who get the flu. Assessing relative fatality depends on considering the *death rates* that are defined by the number of deaths *relative to* the total number of people infected ($\text{fatality}_{\text{flu\_rate}} = .06\%$; $\text{fatality}_{\text{COVID\_rate}} = 4.1\%$). Without these relative comparisons, WNB could result in underestimating the likelihood that any particular infected individual in a population would die from COVID-19. That is, both the number of deaths in isolation and the number of infections in isolation point to the flu being a more severe disease. One cannot compare fatality rates without first dividing deaths by the number of infected individuals.

Importantly, we are taking the stance that to engage in informed decision making, it is important to understand the *case*-fatality rate—the proportion of infected individuals who die—and not simply the proportion of an overall population (e.g., U.S. adults) who die from a given disease. This is important because with a highly contagious disease, such as COVID-19, the absolute number of deaths as a function of the population will increase dramatically as the disease spreads exponentially, whereas the case-fatality rate should remain more stable over time (barring factors such as improved treatment options and better detection/testing availability). For example, in March 2020, early case-fatality rates for COVID-19 were estimated between 1% and 5% across the globe (Yang et al., 2020), yet the media emphasis was on *only* the very low absolute number of deaths in the U.S. at that time (see WHO Situation Reports). This practice of reporting was misleading, as case-fatality rates suggested that if left unchecked, COVID-19 would be responsible for millions of deaths in the U.S. Thus, focusing on the absolute deaths from a virus early in a pandemic could result in underestimating the severity of disease relative to focusing on the case-fatality rate. In the specific context of COVID-19, given the low number of absolute deaths from COVID-19 in mid-March 2020, people may have discounted the virus as a nonsevere health threat. Thus, the purpose of our intervention was to highlight the importance of the COVID-19 case-fatality rate and teach individuals how to correctly calculate and reason about the magnitude of proportions in this critical, real-world context. From a public health perspective, it may be important for people to understand that the overall case-fatality rate of COVID-19 was higher than that of

**Table 1**
*COVID-19 and Flu Statistics as of Mid-March 2020 When Data Were Collected*

| Virus | Number of deaths | Total number of people infected | Case-fatality ratio[a] |
|---|---|---|---|
| Flu | 22,000 | 36,000,000 | 22,000/36,000,000 = 0.00061 = .06% |
| COVID-19 | 9,318 | 227,743 | 9,318/227,743 = 0.041 = 4.1% |

*Note*. Coronavirus disease = COVID-19.
[a] Statistics for the flu are U.S. numbers; statistics for COVID-19 are global numbers, sourced from the Centers for Disease Control and Prevention (2020) and Johns Hopkins University and Medicine (2020) websites in March 2020. As of November 2021, approximately 5 million people had died and nearly 250 million people had been infected by COVID-19. This is a 2.01% case-fatality rate.

the flu, even if any particular individual's risk of death from COVID-19 was not high. Furthermore, the intervention did not attempt to convince individuals that the COVID-19 case-fatality rate represented their own likelihood of dying from COVID, because the risk factors for COVID-19-related death were largely unknown at the time of data collection in March 2020.

In the midst of the pandemic, the denominator, or the number of people who were infected with COVID-19 was unknown. This was especially true in the U.S. given the lack of widespread testing for the virus at the time of data collection in March 2020. Early in the pandemic, the director of the National Institute of Allergy and Infectious Diseases indicated that COVID-19 was likely *10 times* more fatal than the flu (Huang, 2020). Although these public health statements were crucial, because they could arguably influence people to take measures that would literally save lives in the moment as the global pandemic was unfolding, these statements were frequently countered by other contradictory statements from perceived experts and leaders. It is for this reason that we created and tested an educational intervention to decrease adults' WNB errors and to promote accurate interpretation of COVID-19 health statistics. The aim was to test if participants who received the intervention could better recognize the importance of considering the proportion of people who were infected with COVID-19 who died from the virus (i.e., the case-fatality rate). Although there are existing evidence-based interventions that have improved rational number understanding for children (Braithwaite & Siegler, 2020; Fazio et al., 2016; Fuchs et al., 2016; Rittle-Johnson et al., 2001; Schwartz et al., 2011; Sidney et al., 2021), WNB errors continue to persist into adulthood (e.g., DeWolf & Vosniadou, 2015; Obersteiner et al., 2013), and no math cognition interventions for WNB have been developed for adults despite the relevance to health estimation. In the present study, we employed a general instructional intervention, a worked example (McGinn et al., 2015), to target the WNB misconception in adults in the midst of a health crisis.

## Worked Example Intervention

To improve adults' understanding of COVID-19 risk, we included worked examples (Fitzsimmons et al., 2021; McGinn et al., 2015; Renkl, 2014), analogies to familiar contexts (Sidney & Thompson, 2019; Sidney et al., 2021; Thompson & Opfer, 2010; Yu et al., 2020), and number lines to visualize numerical magnitudes (Opfer et al., 2016; Opfer & Siegler, 2007; Opfer & Thompson, 2008; Sidney et al., 2021; Thompson & Opfer, 2008, 2016) which have all been shown to be effective in the domain of math cognition. The brief intervention included a worked example that explained how to calculate flu versus COVID-19 case-fatality rates in a step-by-step manner to increase procedural and conceptual understanding (Rittle-Johnson, 2017). The worked example began with an analogy to a more familiar context, because prior work (Schwartz et al., 2011; Sidney & Alibali, 2017; Sidney & Thompson, 2019) has shown that a preparatory, warm-up exercise prompts people to draw analogies from their relevant prior knowledge to the context at hand. Finally, the worked example concluded with a visualization of flu versus COVID-19 fatality rates on a number line. Number lines, relative to other types of visual models (circle or rectangle area models), have been shown to promote accurate reasoning about the magnitudes of all numbers, including rational numbers (Mielicki et al., 2021; Sidney et al., 2021; Siegler, 2016; Siegler et al., 2011).

That is, adults learned how to think about the holistic magnitude of each virus' rate instead of falling prey to WNB by inaccurately reasoning about the numerator and denominator components in isolation. Decreasing WNB errors should allow individuals to correctly calculate and compare COVID-19 and flu case-fatality rates, which should reinforce the conclusion that COVID-19 is more fatal than the flu and should be taken seriously.

We argue that WNB can impede accurately judging one's disease risk and appropriately calibrating disease worry. Here, we use the term *risk perceptions* to refer specifically to thoughts or feelings about the likelihood of developing a disease, or one's perceived susceptibility. *Disease worry* (Chapman & Coups, 2006; Portnoy et al., 2014) refers to affective concern about developing a disease. Perceived disease risk and disease worry are related, yet distinct, constructs (Ferrer et al., 2016; Taber et al., 2021) that directly influence engagement in health behaviors (Brewer et al., 2007; Hay et al., 2006; Sheeran et al., 2014). We argue that perceived disease risk and worry inherently involve magnitude judgments, as people must consider the magnitude of an outcome (e.g., likelihood of becoming infected or dying from COVID-19). Even if the two disease rates were equal in magnitude, people could be swayed by the size of the whole number components. For example, individuals are less precise when they estimate the magnitude of fractions with larger components relative to smaller components (e.g., 15/30 versus ½; Braithwaite & Siegler, 2018; Fitzsimmons et al., 2020; Woodbury et al., under review). In prior research, constructs similar to WNB—ratio bias and denominator neglect—have been shown to impair health-related decisions (Lipkus & Peters, 2009; Nelson et al., 2008; Peters, 2012; Reyna & Brainerd, 2008; Reyna et al., 2009; Thompson et al., 2021)[1].

## The Present Study

In the present study, we demonstrated that WNB is often the result of focusing on *either* the numerator or denominator in isolation and that focusing on the *relative magnitude* of ratios can help overcome this misconception. That is, we applied theory from math cognition (Siegler et al., 2011) to help adults override the penchant to automatically consider only the whole number components in a ratio instead of processing the ratio's holistic magnitude. Specifically, the integrated theory of whole numbers and fractions development (Siegler, 2016; Siegler et al., 2011) indicates that what all numbers, whole numbers and fractions included, have in common is that their magnitudes can be placed on a number line. It is for this reason that our intervention included a step-by-step explanation of how to calculate case-fatality rates, instead of focusing on independent numerator and/or denominator components, and to represent

[1] We chose to describe this error as whole number bias (e.g., Alibali & Sidney, 2015; Ni & Zhou, 2005; Van Hoof et al., 2015) because it encapsulates a wide range of errors, including ratio bias and denominator neglect. For instance, those who commit denominator neglect pay attention to numerators only, yet evidence from math cognition suggests that people can also ignore numerators and focus on denominators instead. Whole number bias is evident in conceptual understanding of rational numbers (e.g., density: the infinite numbers between any two rational numbers), arithmetic operations (e.g., misapplying whole-number operations during fraction arithmetic), and magnitude knowledge (e.g., comparing numerators or denominators in isolation).

these magnitudes on number lines, so that they could be directly compared to one another.

Our primary hypotheses were that (a) WNB errors, reflected in both strategy reports and performance accuracy on health-related math problems, would be common across conditions at pretest (see Thompson et al., 2020 for example of whole-number bias; WNB errors common early in the pandemic), but that (b) after training, intervention participants would be more accurate on math problems related to COVID-19 and less likely than control participants to report WNB errors in their strategies. In our preregistration (https://osf.io/9hc7d), we stated that we would code participants' strategy reports to evaluate whether they were using math or nonmath strategies and whether their math strategies were correct. Although not stated explicitly in the preregistration, the incorrect strategies that we were most interested in were strategies consistent with WNB (see Thompson et al., 2020).

An additional exploratory aim was to assess the effect of the intervention and math anxiety on risk perceptions, worry, affect, and health behavior. Given our expectation that the intervention would lead to a more accurate understanding of COVID-19 case-fatality rates, we expected that intervention participants, relative to control participants, would report greater perceived risk and worry regarding COVID-19 given their ability to calculate a case-fatality rate. However, this was not explicitly stated in the preregistration. Given our applied interest in promoting engagement in recommended health behaviors, we also explored whether the intervention influenced positive and negative affect and health behaviors (e.g., social distancing, mask wearing) across 10 days following the intervention.

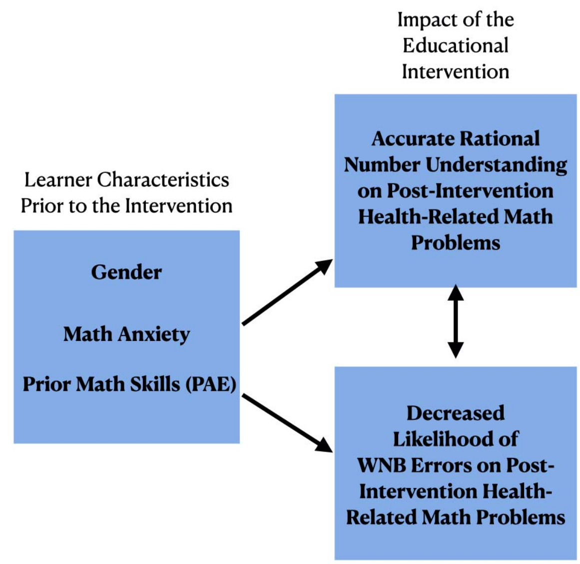
See Figure 1 for our theory of change. We propose that learners bring important individual differences to the learning environment (i.e., general math skills, math anxiety, and gender). The educational intervention's proposed mechanism of change is that it teaches people the importance of computing and comparing the holistic magnitudes of ratios, and indicates that relying on individual whole number components of the ratios leads to errors when thinking about case-fatality rates. Reasoning about holistic magnitudes increases the likelihood of answering the COVID-19 health-related math problems correctly and decreases the likelihood of reporting a WNB strategy to solve these problems. Therefore, we examined whether individual differences shown to be important in other empirical studies of rational number understanding—number line estimation, math anxiety, and gender—also had an effect here.

- First, rational number estimation precision is a proxy for general math skills, because it is strongly correlated with math achievement (Fazio et al., 2014; Siegler & Thompson, 2014; Siegler et al., 2011). People who struggle with rational number magnitudes exhibit higher math anxiety (Sidney, Thalluri, et al., 2019) and score lower on objective measures of numeracy (Choi et al., 2020).

- Second, math anxiety and negative attitudes about math could increase the likelihood of WNB by leading people to avoid thinking deeply about rational numbers. All types of numbers can elicit math anxiety (Ashcraft, 2002; Ashcraft & Krause, 2007), or apprehension about mathematics (Beilock et al., 2010). Math anxiety may be exacerbated for rational numbers, because people report disliking rational numbers more than whole numbers (Sidney et al., 2021).

- Finally, we considered the effects of gender, because women show lower estimation precision (Hutchison et al., 2019; Rivers et al., 2021; Thompson & Opfer, 2008), more math anxiety (Dowker et al., 2016), and more negative math attitudes than men (Sidney et al., 2021).

**Figure 1**
*Theory of Change for Educational Intervention*




*Note.* Note that PAE = percent absolute error, which is a measure of precision on the number line estimation task which taps into underlying magnitude understanding. The mechanism of change involves improving adults' ability to reason more accurately about rational numbers in a variety of health-related math problems after the intervention. This occurs because adults report using fewer whole number bias strategies after engaging with the intervention. See the online article for the color version of this figure.

In the current investigation, we borrowed best practices from interventions with children that could prove worthwhile for dispelling the mathematical misconception, WNB, and thus improve adults' rational number understanding about COVID-19 statistics. To investigate our hypotheses, we collected data from March 24 to April 9, 2020 when COVID-19 "stay-at-home" orders were beginning across the U.S. To the best of our knowledge, no other math interventions have taken a daily diary approach to track the longevity of intervention effects. This study is also unique in that we used a critically important externally valid context to examine theoretically motivated basic science questions about WNB.

At the time of data collection, news and media outlets were providing the kind of statistics presented in our intervention that emphasized overall population case-fatality rates. This information was always presented as ratios, or just as the numerator, and not as percentages, which would have been easier and more intuitive to interpret (Moss & Case, 1999; Siegler et al., 2011). Thus, the goal was to help people interpret the types of statistics that were available through media outlets at the time; the goal was *not* to test whether providing the information in different ways, or providing different types of information (i.e., more personalized), improved comprehension or increased adherence to preventive behaviors. Furthermore, our goal with the present study was to create an intervention that could provide adults with the rational-number skills that they needed to figure out case-fatality rates in any scenario if they were presented with the number of deaths and the total number of infected cases. Thus, rather than targeting the optimal ways that media outlets should report statistics to improve comprehension, we wanted to improve adults' ability to interpret statistics accurately regardless of whether they are presented in an optimal format.

# Method

## Participants

This study was approved by the Kent State University Institutional Review Board; all participants provided online consent for their participation, and their participation was voluntary. Data were collected from March 24 to April 9, 2020 through Qualtrics panels. All recruitment occurred through panels of respondents that were managed by Qualtrics. The research team did not actively recruit for the present study or create any materials to advertise the study. Qualtrics managed the stratification of gender, age, and educational attainment based on agreed-upon quotas. If participants from the Qualtrics panel chose to participate, they clicked on the link for our study which was programmed to run on the Qualtrics platform.

In our preregistration (https://osf.io/9hc7d), we planned to sample 1,200 people and to obtain daily diary data from at least 625. This stopping rule was based on the availability of funds to collect the data via the Qualtrics panel. We sought to recruit equivalent numbers of males and females and to stratify by educational attainment to represent education levels in the U.S. population, although in the final sample, a smaller proportion of people (3.93%) had earned less than a high school diploma than in the overall U.S. population (12.3%, United States Census Bureau, 2018). Approximately 75% of participants self-identified as White, 46% identified as male, 41% reported being employed for wages, and 70% reported having between some college experience and a graduate degree. The average reported age of participants was 46.90 years ($SD$ = 17.34 years; range: 18–85 years). See Table 2 for demographics.

We embedded checks on attention and engagement during the baseline assessment and over the 10 days of daily diaries (Behrend et al., 2011; Hauser & Schwarz, 2016; Meyerson & Tryon, 2003). Preregistered data cleaning processes (e.g., nonsensical open-ended responses, patterned responding, use of less than 10% of the line on estimation tasks, etc.), detailed in Appendix A, reduced the sample from an initial 2,693 adults who consented to 1,297 who provided adequate data and compliance. Participants were primarily excluded based on short completion time or failing one of the two attention checks. There were some differences between those included and those excluded from analysis: those excluded were younger and more likely to be white, female, students, or self-employed. They were also more likely to report lower income, to have taken fewer math courses, to be incorrect on the objective numeracy and baseline health-related math problem-solving question, and less likely to be retired or employed for wages.

**Table 2**
*Participant Characteristics*

| Sociodemographic factors | $N$ | % |
|---|---|---|
| Race and ethnicity | | |
| White | 969 | 74.71 |
| Black or African American | 134 | 10.33 |
| American Indian or Alaska Native | 8 | 0.62 |
| Asian | 50 | 3.86 |
| Native Hawaiian or Pacific Islander | 3 | 0.23 |
| Hispanic or Latino | 51 | 3.93 |
| Other | 13 | 1.00 |
| Did not report | 3 | 0.23 |
| Multiple | 66 | 5.09 |
| Employment | | |
| Employed for wages | 536 | 41.33 |
| Self-employed | 112 | 8.64 |
| Out of work >1 year | 77 | 5.94 |
| Out of work <1 year | 55 | 4.24 |
| Homemaker | 91 | 7.02 |
| Student | 96 | 7.40 |
| Retired | 315 | 24.29 |
| Rather not report | 15 | 1.16 |
| Gender | | |
| Male | 593 | 45.72 |
| Women | 697 | 53.74 |
| Other gender | 2 | <0.01 |
| Rather not report | 5 | <0.01 |
| Education | | |
| <High school and high school | 389 | 29.99 |
| Some college or associates | 472 | 36.39 |
| Bachelor's | 285 | 21.97 |
| Graduate degree | 151 | 11.64 |
| Income | | |
| <15,000 | 155 | 11.95 |
| 15,000–24,999 | 147 | 11.33 |
| 25,000–34,999 | 146 | 11.26 |
| 35,000–49,999 | 190 | 14.65 |
| 50,000–74,999 | 252 | 19.43 |
| 75,000–99,999 | 154 | 11.87 |
| 100,000–149,999 | 140 | 10.79 |
| 150,000–199,999 | 41 | 3.16 |
| >200,000 | 33 | 2.54 |
| Would rather not report | 39 | 3.01 |
| Math courses taken | $M$ ($SD$) | Range |
| Sum of math courses (max of 11) | 3.84 (2.59) | 0–11 |

*Note*. For the purpose of analyses, males were coded as 1, and participants who selected "female," "other," or chose not to answer were coded as 0. For ease of interpretation, we refer to the latter group as "women," although we acknowledge this is an inaccurate category label.

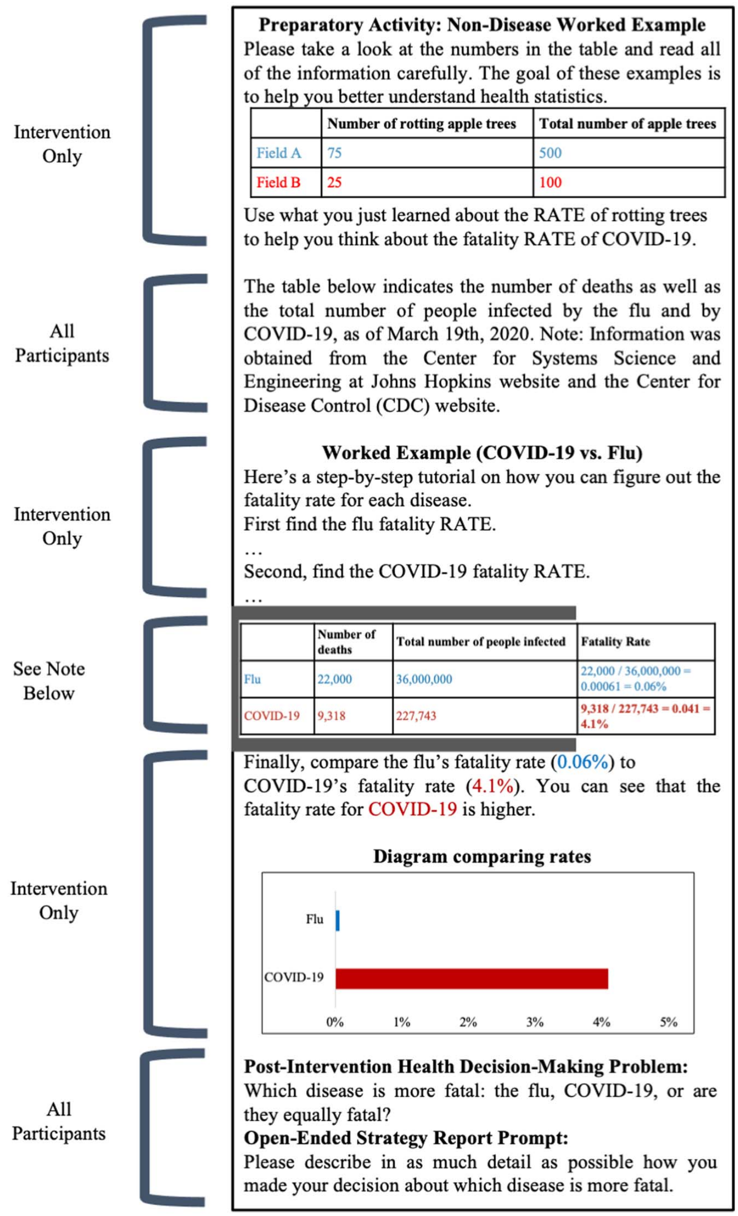
The analytic sample including daily diary data was reduced further, because only 709 participants completed at least some of the daily diaries, and other participants were excluded based on compliance or failure to respond to attention checks within the diaries (see Appendix A). The final diary sample was 627 individuals who completed 75% (4,703/6,270 = 0.75) of the possible diary signals ($M$ = 7.58, $SD$ = 3.11).

## Experimental Design and Procedure

A flowchart of the study procedures can be found in Appendix B, and a comparison of the intervention and control condition is illustrated in Figure 2. The full educational intervention is available on OSF (https://osf.io/45ycd/).

**Figure 2**

*Comparisons Between the Educational Intervention and the Control Condition*




*Note*. The coronavirus disease (COVID-19) and flu deaths and infections contained within the gray background were presented to control participants in a static 2 × 2 contingency table in black font. Intervention participants were shown how to calculate the fatality rates in the tables as the information in blue and red font unfolded in a dynamic way throughout the worked example. Whereas we have illustrated the full information that control participants received, the full text that the intervention participants received is in the Supplemental Materials. See the online article for the color version of this figure.

We randomly assigned participants to either an educational intervention or a business-as-usual control condition. Random assignment was managed by the Qualtrics platform and programmed to keep the numbers of participants assigned to each condition fairly equal as data collection progressed. No feedback on participants' performance was provided at any point during the study, and participants progressed through the study at their own pace.

Participants completed a baseline survey and 10 days of follow-up daily diary surveys. On day 1, participants completed an approximately 40-min online study consisting of a baseline survey, randomly assigned intervention condition, and an immediate post-test. The baseline survey assessed (a) sociodemographic factors, (b), perceived knowledge of COVID-19 and positive and negative affect, (c) math tasks known to predict WNB, and (d) a pretest health decision-making problem. Participants who were randomly assigned to the *educational intervention condition* completed an intervention that taught them how to consider the relation between numerators and denominators when reasoning about rates and how to calculate case-fatality rates for different diseases. In the *business-as-usual control condition*, participants saw relevant statistics (i.e., number of deaths and number of infected individuals), but were not shown how to calculate case-fatality rates. After the intervention, participants completed three health-related math problems. After solving each health-related math problem, participants reported the strategy that they used to solve each problem (see Appendix C). Participants then indicated their risk perceptions and worry about COVID-19, their positive and negative affect, need for cognition, and health literacy. For the next 10 days, participants completed experience sampling via a brief online survey at the same time every evening in an effort to track the impact of the educational intervention on their risk perceptions, worry, positive and negative affect, and preventive health behaviors.

## Materials

Measures are described in the order administered. Reliability, means, and standard deviations for all measures, and correlations among measures, are reported in Table 4.

### *Sociodemographic Factors*

Prior to giving informed consent, participants indicated eligibility by confirming that they were 18+ years of age and held U.S. citizenship. After providing informed consent, participants answered sociodemographic questions, including age, gender, employment status, level of educational attainment, number and type of math courses taken in high school and college, the type of device they used to complete the survey, race/ethnicity, household income, and zip code. Responses of “Would rather not report” for the employment variable were coded as missing. Race was coded as white versus nonwhite for statistical analyses.

### *Baseline Health-Related Math Problem Solving*

Participants determined which of two hypothetical diseases was more fatal in a single pretest problem. Disease A included a bigger numerator and a bigger denominator (analogous to flu statistics); Disease B included a smaller numerator and a smaller denominator (analogous to COVID-19 statistics). The numbers chosen in the baseline problem for Disease A were analogous to the numbers for COVID-19 in China as of February 2020, and the numbers for Disease B were created by multiplying the total cases by 30 and creating a fatality rate that was 1% of that. We chose these numbers, because they led to a difference in fatality rate that was comparable to the post-test Problem 1 (i.e., COVID-19 versus the flu).

### *Baseline Knowledge of COVID-19 and Emotion*

**Perceived Knowledge of COVID-19.** After answering the baseline health-related math problem, participants read a brief paragraph about COVID-19:

> Many questions in this survey will be about the novel coronavirus, COVID-19. Coronavirus disease 2019 (COVID-19) is a respiratory illness that can spread from person-to-person. The virus that causes COVID-19 is a novel coronavirus that was first identified during an investigation into an outbreak in Wuhan, China. The first case of COVID-19 in the United States was reported on January 21, 2020.

After participants read this paragraph, they were asked, Overall, how would you rate your level of knowledge about COVID-19 (for example, what it is, how it is transmitted, how to protect yourself, etc.)? Their answers ranged from 1 = *No knowledge at all* to 4 = *A lot of knowledge*. Participants could also choose “Do not know.”

Because we collected these data in March 2020 at the beginning of the COVID-19 pandemic, we included a measure of perceived COVID-19 knowledge, as variability in pre-existing COVID-19 knowledge could influence participants’ likelihood of endorsing that COVID-19 was more fatal than the flu.

**Negative and Positive Affect.** Participants rated their current affective experience for five specific positively valenced (relief, amusement, affection, happiness, and interest) and six negatively valenced (disgust, sadness, fear, guilt, distress, and anger) emotion words on a Likert scale (1 = *None* to 7 = *Strong*; Coifman et al., 2016). We created separate aggregate scores to index negative and positive emotions.

### *Baseline Math Tasks*

The order of math attitudes and math anxiety measures was randomized. These measures were presented before the math skills measures because in past research (Sidney et al., 2021) participants reported more negative attitudes if they completed these math tasks before rating their attitudes. The order of the math skills measures was randomized for all participants. Unrelated to the present study’s hypotheses, we also collected data on several additional math measures: fraction equivalence (Fitzsimmons et al., 2020), subjective numeracy (Fagerlin et al., 2007), and objective numeracy (Cokely et al., 2012). Subjective numeracy and objective numeracy measures were assessed because they are typically used in the health decision-making literature as measures of preferences for math and computational ability.

In a recently published study from our lab (Thompson et al., in press), measures of subjective (Fagerlin et al., 2007) and objective numeracy (Weller et al., 2013) from the health decision-making literature were not significant predictors of health decision-making accuracy when entered simultaneously into regression models with measures of magnitude understanding from the math cognition literature (i.e., magnitude comparison accuracy and multistep fraction arithmetic performance). Therefore, we did not include subjective and objective numeracy as predictors in our logistic regression models, but we did confirm that performance on these measures did not differ by experimental condition to ensure that random assignment to condition was successful.

In addition, our previous research (Fitzsimmons et al., 2020) indicated that adults’ WNB errors on a fraction estimation task were correlated with fraction equivalence knowledge—as measured by asking participants to decide whether sets of fractions were equivalent in magnitude. Therefore, we thought that it was necessary to also ensure that our experimental groups in the present study did not differ on equivalence accuracy.

**Math Anxiety.** Participants rated their overall math anxiety (Ashcraft, 2002) and math anxiety about specific types of numbers—whole numbers, fractions, percentages and whole number frequencies—for five items on a scale ranging from 1 = *Not anxious* to 10 = *Very anxious*. We calculated an average math anxiety score across items.

**Math Attitudes Questionnaire.** Participants answered 20 questions (Sidney et al., 2021) pertaining to their attitudes about math in general as well as their specific attitudes about whole numbers, fractions, and percentages. This math attitudes questionnaire involves subscales for self-perceived ability, preferences, frequency of use, and importance. We calculated an average math attitude score across items.

**Math Skills.** Math skills were measured with number line estimation tasks for fractions, whole number frequencies, and percentages. Number line estimation is quick (Fazio et al., 2016) and easy to administer online (e.g., with Qualtrics sliders) or on paper.[2] In three blocks separated by number type, participants made estimates either for one fraction at a time (estimated on number lines with 0 as the left endpoint and 5 as the right endpoint: Sidney et al., 2021; Siegler et al., 2011), one whole number frequency at a time (estimated on number lines with 0 out of 100 as the left endpoint and 100 out of 100 as the right endpoint), or one percentage at a time (estimated on number lines with 0% as the left endpoint and 5% as the right endpoint). See Figure S1 for an example trial from each range. Participants completed 9 or 10 trials in each block. Participants' performance on each trial was measured as percent absolute error (PAE; Siegler & Booth, 2004): PAE = [(|estimate − true value|)/numerical range] × 100. PAE was averaged across all trials with higher scores indicating greater error of estimation, or worse performance.

The rationale for presenting percentage estimation questions in the 0%–5% range is that it corresponded to the 0–5 fraction estimation range, and the COVID-19 case-fatality rate at the time of data collection was less than 5%. To the best of our knowledge, this is the first study to ask participants to estimate whole number frequencies in the 0 out of 100 to 100 out of 100 range. While it is true that adults are very accurate when estimating numbers in the 0–100 range (Siegler & Opfer, 2003), this does not guarantee that they will accurately estimate ratios in the 0 out of 100 to 100 out of 100 range. To foreshadow our results, all three estimation scales were internally reliable and correlated strongly with one another, indicating that they tapped the same construct of magnitude understanding (see Table S1).

### *Educational Intervention Versus Business-As-Usual Control*

Participants were randomly assigned to one of the two experimental conditions: the educational intervention to combat WNB or the business-as-usual control condition.

### *Educational Intervention Condition*

The main goal of our intervention was to facilitate adults' procedural and conceptual understanding of *how* and *why* they should calculate and compare case-fatality rates. As noted in Table 1, participants were taught about case-fatality rates using the numbers of COVID-19 deaths and cases worldwide in mid-March 2020 when the data were collected. Information about the pandemic is ever-evolving, as COVID-19 health statistics update daily. Our intervention was developed to help people interpret the dynamic nature of these statistics. That is, the intervention provided participants with procedural and conceptual knowledge about how to calculate and compare case-fatality rates, such that they would be able to transfer this math skill to future-related problems or health statistics, including those outside of the COVID-19 context.

First, participants in the intervention condition were instructed to complete a diagram-based preparatory activity (Schwartz et al., 2011; Sidney & Thompson, 2019) in which we visually illustrated that some numeric comparisons require taking both numerators and denominators into account by calculating a ratio. This preparatory activity, in a more accessible, concrete context (e.g., two orchards with differing rates of rotting apple trees; see Figure 2), was designed to facilitate drawing analogies to the COVID-19 context.

Then, participants studied the focal worked example, which was designed to decrease WNB errors (i.e., directly comparing numbers of deaths or infections) by teaching participants how to consider the relation between the number of deaths and number of cases when calculating the case-fatality rate with given frequency information. Participants saw the number of deaths (numerator) and total number infected (denominator) for both the flu and COVID-19 in a 2 × 2 contingency table (see Table 1). Participants were instructed how to calculate and compare the COVID-19 versus flu fatality rates via worked example, demonstrating each step in turn (McGinn et al., 2015) and ending with a number line comparing the magnitudes of the case-fatality rates for COVID-19 and the flu (Hamdan & Gunderson, 2017; Sidney, Thompson, et al., 2019). In summary, our intervention leveraged empirical findings from cognitive science to help adults avoid WNB errors.

Worked examples are a well-researched intervention for demonstrating mathematical procedures in a step-by-step manner, such that students can learn the correct procedures that will help them overcome common mathematical misconceptions. At each step of the problem, students learn information (or are asked to generate self-explanations) about why these are the correct or incorrect steps to solve the problem. The worked example taught participants the steps necessary to solve all of the post-test health-related math problems (described above) that differed in surface-level features (e.g., numbers, ratios, case-fatality rates, infection rates, and contingency tables versus word problems) and scenarios (COVID-19 versus flu, COVID-19 across time, and infection rates).

### *Business-As-Usual Control Condition*

Control participants saw the same 2 × 2 contingency table as the educational intervention group and were asked to solve the same problem (i.e., "Which disease is more fatal?"). See Figure 2 and Table 3 for the exact values shown. However, the control

[2] Researchers interested in assessing whole number bias in number-line estimation tasks should consult Braithwaite and Siegler (2018) and Fitzsimmons et al. (2020). Detailed discussions of estimation precision calculations (PAE) can be found in Siegler et al. (2011).

**Table 3**
*Measures of Health-Related Math Problem Solving*

**Pretest**

| | Total # deaths | Total # infected |
|---|---|---|
| A | 2,125 | 55,924 |
| B | 16,777 | 1,677,720 |

Which disease is more fatal: *Disease A*, Disease B, or are they equally fatal?
Please describe in as much detail as possible how you made your decision about which disease is more fatal.

**Postintervention Problem 1**

| | Total # deaths | Total # infected |
|---|---|---|
| Flu | 22,000 | 36,000,000 |
| COVID-19 | 9,318 | 227,743 |

Which disease is more fatal: the flu, *COVID-19*, or are they equally fatal?
Please describe in as much detail as possible how you made your decision about which disease is more fatal.

**Postintervention Problem 2**

| | Total # deaths | Total # infected |
|---|---|---|
| Jan. 30 | 170 | 7,818 |
| Feb. 6 | 565 | 28,276 |

Did the COVID-19 fatality rate: increase, *decrease*, or stay the same from January 30th to February 6th?
Please describe in as much detail as possible how you made your decision about the possible change in fatality rates.

*Note*. COVID-19 = coronavirus disease. All health-related math problems were forced choice. We have italicized the correct responses here for readers. All strategy reports were open-ended. Importantly, the wording of the pretest, postintervention Problem 1, and postintervention Problem 2 health-related math problems was *identical* for all participants, regardless of experimental condition.

participants were not told which numbers to attend to, because they did not engage in the educational intervention.

## Postintervention Health-Related Math Problem Solving

Immediately after the intervention, participants completed health-related math problems, and accuracy was assessed for each problem with a forced choice question (e.g., Which disease is more fatal?). There were three possible responses; thus, chance accuracy on each problem was 33.33%. See Appendix E for full statistical models.

Problem 1 assessed whether participants correctly answered a question about case-fatality rates for COVID-19 versus the flu. We expected that those in the intervention condition would correctly respond that COVID-19 was more fatal than the flu, given that this was the topic of the intervention. However, if participants selected flu as the more fatal illness, then this response would be consistent with WNB. We anticipated that this would be a common response for those assigned to the control condition, given that they did not engage with the intervention, or for those individuals in the experimental condition if they did not "uptake" the lesson.

In Problem 2, participants compared COVID-19 case-fatality rates across time. This item was designed to be misleading as a *strong test* of participants' ability to apply the mathematical computations introduced in the intervention because the correct answer was inconsistent with WNB and media coverage. We specifically chose to present a time period in which the case-fatality rate for COVID-19 *decreased*: January 30th vs. February 6th (based on statistics from the WHO Situation Reports). We suspected that participants would anticipate that the COVID-19 case-fatality rate would increase over time due to (a) WNB—because the absolute numbers of COVID-19 cases and deaths *did* generally increase across time as the pandemic unfolded—or (b) increased media coverage conveying that, indeed, the fatality rate was increasing over time (see COVIDView weekly updates launched on the Centers for Disease Control and Prevention (CDC) website on March 28th, the week that we began collecting data: https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covidview/past-reports/04032020.html). Thus, if participants reasoned that case-fatality rates increased over time, they would get this problem wrong, and their response would be consistent with WNB. In contrast, if the intervention had successfully ameliorated WNB errors and strategies, then participants would be more likely to respond correctly that the case-fatality rate had decreased over time, and they should be less likely to report WNB strategies in their open-ended strategy reports. If participants who were exposed to the educational intervention outperformed those who were randomly assigned to the control condition on this problem and related strategy report, this would be strong evidence that the intervention successfully taught participants how to consider the holistic magnitude of rational numbers when calculating case-fatality rates to solve a novel problem.

For Problem 3, participants compared infection rates in Italy and China. News sources noted that Italy's infection rate, death toll, and fatality rate were especially high (e.g., Di Donato et al., 2020) and that Italy was the pandemic's new epicenter (Horowitz et al., 2020) being hit harder than China (Perez-Pena, 2020). It is possible that attention in the press led many participants, even those in the control group, to answer this problem correctly (about 60% of participants across both conditions). Although 60% may not seem like a large proportion, approximately 45% of participants who missed the pretest problem correctly answered Problem 3. However, only 14% of participants who missed the pretest problem correctly answered postintervention Problem 2. Thus, it seems that Problem 3 was likely too easy because it could be solved with nonmath information. In fact, 6% of participants mentioned news and media in their strategy reports for this problem. Interested readers can find our analyses of this problem in the Supplement.

## Strategy Reports and Coding

In addition to decision-making response choice, we collected strategy reports, a form of converging evidence that could also reveal WNB errors. This converging evidence is especially critical, because the health-related math problem format was adapted from Waters et al. (2007) in which there were only three possible responses. Thus, guessing correctly by chance could occur

33.33% of the time. Participants provided immediate retrospective strategy reports for each of the health-related math problems by explaining their responses in a text box in response to the prompt, "Please describe in as much detail as possible how you made your decision about . . . " (Fazio et al., 2016; Fitzsimmons et al., 2020; Sidney, Thalluri, et al., 2019; Siegler et al., 2011; Siegler & Thompson, 2014; Thompson et al., in press). Strategy reports can reflect the impact of experimental condition on health-related math accuracy by allowing insight into the specific correct or incorrect methods individuals use to solve math problems (e.g., Siegler & Thompson, 2014). In sum, we used strategy reports as converging evidence that participants' answers on the health-related math problems were indicative of committing WNB errors.

The coding scheme included three codes to assess instances of WNB error: numerator, denominator, and larger numbers (see Appendix C for the full list of codes and descriptions). Note that participants explicitly stated using all three WNB substrategies. Numerator referred to mentioning just the number of deaths, denominator referred to mentioning just the number of cases, and larger numbers referred to a more general reference to the size of numbers without reference to deaths or cases specifically. These codes reflected mentioning either the numerator or the denominator in isolation when making decisions (e.g., *Look at the biggest number; How many died from it; Disease B has more people infected*). For analyses, we collapsed across these three WNB codes. Participants who received one of the three WNB error codes *but also* reported a more mathematically sophisticated strategy (i.e., indicated calculating a rate, including transformations to easier to handle numbers, such as decimals or percentages) were not included as having made WNB errors in the analyses reported below. Very few participants reported both rate and WNB strategies across the health-related math problems: pretest: $n = 14$ (1%), postintervention Problem 1: $n = 9$ (1%), and postintervention Problem 2: $n = 5$ (0%).

Two coders (Marta Mielicki and Erika Schemmel) independently coded responses from 30 participants (for a total of 120 responses); interrater reliability was high (97% interrater agreement across all codes and responses). The two raters resolved their disagreements through a conversation. Marta Mielicki then coded 13.3% and Erika Schemmel coded 87.7% of the remaining strategy reports. The primary strategy coder was not privy to the full experimental design and hypotheses. Both the primary and secondary coders were blind to condition, as they coded strategy reports and whether the participant got the answer right or wrong.

## Postintervention Affect and Health Cognitions

### *Negative and Positive Affect*

This was the same measure administered at baseline.

### *Risk Perceptions*

*Risk perceptions* (adapted from Klein & Ferrer, 2018) were assessed as the perceived likelihood that oneself and close others would be infected with COVID-19 in the next year. Items assessing risk perceptions for oneself assessed absolute cognitive ("Overall, how likely is it that *you* will be infected with COVID-19 in the next year?"; 1 = *not at all likely* to 5 = *extremely likely*; 6 = *do not know*), comparative ("Overall, how do you think *your* chance of being infected with COVID-19 in the next year compares to other women or men of your age in the United States?"; 1 = *much less likely* to 5 = *much more likely*; 6 = *do not know*), and experiential risk ("I feel very vulnerable to being infected with COVID-19 in the next year"; 1 = *strongly disagree* to 5 = *strongly agree*; 6 = *do not know*). *Risk perceptions* for others were assessed with the same three items and instead referenced "family members or friends." "Do not know" responses were coded as missing data. Risk perceptions for oneself and others were correlated ($r = .69$, $p < .0001$); thus, all six items were averaged to form one scale.

### *COVID-19 Worry*

*Worry* (adapted from Taber et al., 2019; Weinstein et al., 2007) was assessed as the extent of worry that oneself and close others would be infected with COVID-19. Worry for oneself was assessed with two items: "How much do *you* worry about being infected with COVID-19?," and "How anxious are *you* about being infected with COVID-19?" (1 = *not at all* to 4 = *a lot*). *Worry for others* was assessed with the same two items, but referenced "family members or friends." Worry for oneself and for others were correlated ($r = .76$, $p < .0001$); thus, all four items were averaged to form one scale.

## Postintervention Individual Differences Assessment

These individual differences measures were administered in the same order for all participants at the end of the study. We did not expect responses to be influenced by exposure to the intervention condition. Need for cognition and health literacy measures were included in a subset of analyses. However, these variables did not predict the likelihood of correctly answering health-related math problem-solving accuracy or reports of WNB aligned strategies. Health literacy was negatively related to the proxy for overall math ability (PAE) and need for cognition. Higher need for cognition was related to lower math anxiety and higher, more positive math attitudes. Finally, perceived COVID-19 knowledge was related to higher, more positive math attitudes, higher health literacy, and higher need for cognition (see Table 3). We also included a trait anxiety measure that was not central to the current hypotheses and thus is not discussed further.

### *Need for Cognition*

This scale consists of the average of six items assessing the extent to which people enjoy engaging in the process of thinking (Cacioppo & Petty, 1982; Coelho et al., 2018). This was relevant to the present study given that computing and comparing ratios are cognitively demanding tasks. A sample item includes: *I would rather do something that requires little thought than something that is sure to challenge my thinking abilities.* Participants rated themselves on a scale from 1 = *Extremely Uncharacteristic* to 5 = *Extremely Characteristic*. We included this measure to assess whether people sought out experiences that allowed them to engage in the process of thinking because these people may be more likely to engage with the numerical information in the health-related math problems.

### Health Literacy

The health literacy measure (Chew et al., 2008) included three items (i.e., confidence in filling out medical forms on own, have someone help read health materials, or have problems learning about a medical condition because of a difficulty understanding written materials). Participants rated themselves on a scale of 1 = *None of the time/Not at all* to 5 = *All of the time/Extremely*. A health literacy score was computed as the sum of responses. We included this measure to assess whether participants differed in their confidence and facility with understanding health-related information and whether this was associated with health-related math problem solving.

### Daily Diaries (Days 2 Through 11)

For the 10 days following the baseline assessment and intervention, participants were prompted via email each evening to answer questions about their emotional experiences, behaviors enacted, and perceptions of risk and worry related to COVID-19. Each of the indices demonstrated excellent measurement sensitivity both within ($R_C$) and between ($R_{KF}$) persons over the 10-day assessment period (Cranford et al., 2006). See Table 4.

### Negative and Positive Affect

This was the same measure administered at baseline and postintervention.

### Preventive Health Behaviors

Participants reported the frequency of 13 behaviors (e.g., social distancing, hand washing, wearing masks, and avoiding group activities; Appendix D) recommended broadly by public health experts and agencies (e.g., CDC). We summed the number of recommended behaviors participants reported engaging in at each diary signal, so that each participant had a daily score of behaviors, 0–13.

### Risk Perceptions

*Risk perceptions* (adapted from Klein & Ferrer, 2018) were assessed with the two absolute cognitive risk items (i.e., risk for oneself and others) included at baseline (comparative and experiential risk items were not included due to the need for abbreviated measures in daily diaries). These two items were highly correlated ($r = .83$, $p < .0001$) and were averaged.

### COVID-19 Worry

*Worry* (adapted from Weinstein et al., 2007) was assessed with the items from baseline concerning how much individuals worry about themselves and their family members and friends being infected with COVID-19. These two items were highly correlated ($r = .80$, $p < .0001$) and were averaged.

## Overview of Analyses

The data analysis plan was preregistered on OSF before data collection began. Analyses involving variables assessed in the baseline session were conducted in *R* Version 3.6.0 (R Core Team, 2019). First, we tested whether random assignment of participants to condition was successful in yielding equivalent groups at baseline for key variables. Analyses indicated that only gender was not equally distributed across conditions; thus, we controlled for gender in all subsequent analyses. Second, we conducted a series of logistic regressions testing the effects of condition (i.e., educational intervention versus business-as-usual control) on each of the postintervention health-related math problems both in (a) simplified models controlling for gender alone and in (b) "full" models. The full models controlled for gender, additional sociodemographic factors of age, race, educational attainment, and expected predictors of rational number decision-making accuracy, that is, magnitude knowledge (a composite score for PAE across three number line estimation scales), baseline health-related math problem-solving accuracy, math anxiety, math attitudes, health literacy, need for cognition, and perceived COVID-19 knowledge. The majority of these covariates were specified in the preregistration and were included to improve power by explaining greater variance in decision-making accuracy. We also covaried age and race because both factors have been associated with perceived disease risk (e.g., Taber et al., 2017). Third, we conducted logistic regressions testing the effects of condition on WNB strategy reports for the health-related math problems assessed postintervention, controlling for gender. Fourth, we conducted a series of linear regressions testing the effects of condition on COVID-19 risk perceptions, worry, and positive and negative affect immediately following the intervention while controlling for gender. We report the final fixed effects, including 95% CIs, for our results below.

Finally, we explored the effect of condition on daily risk perceptions, worry, affect, and health behaviors reported during the 10 days of daily diaries by applying multilevel modeling (via SAS 9.4 Proc Mixed). See the Supplement for details about this analysis.

# Results

## Confirming Random Assignment and Baseline Rates of WNB

We assessed differences across conditions to confirm successful random assignment. There were no condition differences in the baseline health-related math problem, age, education, employment status, number of math courses taken, race/ethnicity, number line estimation PAE, equivalence accuracy, math anxiety, math attitudes, subjective numeracy, objective numeracy, need for cognition, health literacy, baseline negative or positive affect, or the number of participants who failed attention checks (all $p > .09$). There were more males randomly assigned to the control group (52.78%) than the intervention (47.2%), $\chi^2(1) = 4.47$, $p = .034$. As per our preregistration analysis plan, we included gender as a covariate in subsequent analyses. Only 43.10% of participants correctly indicated that hypothetical Disease A was more fatal than Disease B in the pretest health-related math problem, and this did not differ across conditions, control: 42.99%; intervention: 43.21%, $\chi^2(1) < .01$, $p = .982$. Of those who were incorrect, 43.72% (control: 42.68%; intervention: 44.73%) said that Disease B was more fatal, consistent with WNB. The percentage of participants who reported WNB in their strategy reports was approximately equal in the control (27.57%) and experimental (28.70%) groups, $\chi^2(1) =$

**Table 4**
*Descriptive Statistics and Correlations for Measures in Survey*

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Percent absolute error | — | | | | | | | | | | | | | | | | | | |
| 2. Need for cognition | **−0.17** | — | | | | | | | | | | | | | | | | | |
| 3. Math anxiety | **0.37** | **−0.20** | — | | | | | | | | | | | | | | | | |
| 4. Math attitudes | **−0.34** | **0.44** | **−0.35** | — | | | | | | | | | | | | | | | |
| 5. Covid knowledge | 0.01 | **0.14** | 0.01 | **0.19** | — | | | | | | | | | | | | | | |
| 6. Health literacy | **−0.18** | **0.14** | **−0.28** | 0.01 | **0.08** | — | | | | | | | | | | | | | |
| 7. Baseline risk perceptions | *−0.06* | −0.02 | **0.15** | *0.05* | **0.10** | **−0.11** | — | | | | | | | | | | | | |
| 8. Baseline perceived worry | **0.07** | −0.02 | **0.25** | *0.05* | **0.16** | **−0.17** | **0.59** | — | | | | | | | | | | | |
| 9. Perceived severity | 0.04 | 0.03 | **0.06** | −0.02 | **0.12** | **0.10** | **0.24** | **0.36** | — | | | | | | | | | | |
| 10. Pretest positive affect | **0.14** | **0.10** | 0.03 | **0.17** | **0.10** | **−0.18** | **−0.06** | **−0.13** | **−0.10** | — | | | | | | | | | |
| 11. Pretest negative affect | **0.07** | −0.05 | **0.32** | **0.07** | **0.09** | **−0.22** | **0.36** | **0.53** | **0.15** | **−0.07** | — | | | | | | | | |
| 12. Post-test positive affect | **0.12** | **0.12** | 0.03 | **0.21** | **0.07** | **−0.17** | **−0.10** | **−0.13** | **−0.10** | **0.77** | **−0.06** | — | | | | | | | |
| 13. Post-test negative affect | **0.10** | **−0.07** | **0.35** | 0.02 | **0.06** | **−0.23** | **0.36** | **0.53** | **0.17** | **−0.08** | **0.84** | **−0.08** | — | | | | | | |
| 14. Daily diary positive affect | **0.08** | **0.16** | **−0.02** | **0.17** | 0.04 | *−0.07* | **−0.24** | **−0.27** | **−0.12** | **0.63** | **−0.20** | **0.71** | **−0.19** | — | | | | | |
| 15. Daily diary negative affect | **0.11** | −0.03 | **0.25** | −0.01 | **0.10** | **−0.26** | **0.37** | **0.51** | **0.17** | **−0.11** | **0.70** | **−0.11** | **0.72** | **−0.27** | — | | | | |
| 16. Daily diary risk perceptions | 0.01 | −0.01 | **0.12** | −0.02 | **0.09** | **−0.15** | **0.67** | **0.45** | **0.17** | −0.05 | **0.35** | *−0.08* | **0.34** | **−0.19** | **0.43** | — | | | |
| 17. Daily diary perceived worry | **0.13** | −0.05 | **0.20** | −0.05 | **0.16** | **−0.14** | **0.63** | **0.78** | **0.38** | **−0.14** | **0.52** | **−0.13** | **0.52** | **−0.28** | **0.57** | **0.63** | — | | |
| 18. Daily diary health behaviors | **0.22** | **0.16** | **0.20** | **0.15** | **0.22** | **−0.20** | **0.21** | **0.35** | **0.21** | **0.13** | **0.32** | **0.15** | **0.34** | **0.13** | **0.33** | **0.12** | **0.35** | — | |
| 19. Age | −0.01 | **−0.10** | **−0.07** | **−0.15** | *−0.05* | **0.28** | −0.04 | *−0.05* | **0.17** | −0.03 | **−0.09** | **−0.06** | **−0.06** | −0.04 | **−0.08** | **−0.10** | −0.06 | **−0.15** | — |
| *M* | 16.67% | 3.39 | 4.59 | 4.08 | 3.41 | 9.78 | 2.71 | 2.57 | 3.57 | 3.32 | 3.08 | 3.07 | 2.89 | 2.22 | 1.92 | 5.57 | 5.36 | 6.90 | 46.90 |
| *SD* | 8.68% | 0.81 | 2.52 | 1.04 | 0.66 | 2.85 | 0.89 | 0.88 | 0.58 | 1.30 | 1.51 | 1.37 | 1.59 | 0.81 | 0.90 | 2.21 | 1.73 | 2.11 | 17.34 |
| Reliability | 0.89 | 0.79 | 0.92 | 0.97 | – | 0.7 | 0.88 | 0.92 | 0.87 | 0.76 | 0.86 | 0.8 | 0.89 | $R_{KF}$ = .99; $R_C$ = .76 | $R_{KF}$ = .99; $R_C$ = .58 | $R_{KF}$ = .99; $R_C$ = .70 | $R_{KF}$ = .99; $R_C$ = .71 | $R_{KF}$ = .97; $R_C$ = .38 | |
| *N* | 1297 | 1272 | 1297 | 1296 | 1280 | 1297 | 1086 | 1274 | 1283 | 1295 | 1296 | 1290 | 1290 | 627 | 627 | 585 | 627 | 627 | 1297 |

*Note*. COVID-19 = coronavirus disease. All reliability measures are alphas except for the daily measures. Between-person reliability is denoted as $R_{KF}$, Estimating measurement precision regarding systematic between-person differences across days); within-person reliability is denoted as $R_C$; estimating measurement precision regarding systematic change of person from day to day (Cranford et al., 2006). Bolded values are significant at $p < .05$. Italicized values are marginally significant at $p < .1$.

0.15, $p = .695$. Furthermore, consistent with our hypotheses, many participants mentioned WNB errors in their strategy reports (see Appendix F).[3]

## Immediate Effect of the Intervention on Health-Related Math Accuracy

### *Postintervention Health-Related Math Problems*

First, consistent with hypotheses, the intervention relative to control group performed more accurately on health-related math Problem 1, in which participants compared rates of COVID-19 and the flu (Table 5; recall that chance performance on each problem was 33.33% because there were three possible answers); 84% of participants (550 of 655) assigned to the intervention correctly answered this problem compared to 60% of participants (385 of 642) assigned to the control group.[4] This effect held when conducting multivariate analyses controlling for gender and for additional covariates typically associated with accuracy on math tasks (see Appendix Table E1 for full model).

In addition, participants were more likely to choose the specific incorrect response that reflected WNB. Of those participants who were incorrect across the two groups, a larger proportion (67.13%) chose the response consistent with WNB (flu was more fatal) compared to the proportion who chose the incorrect response that was not consistent with WNB (the diseases were equally fatal; 32.87%), $\chi^2(1) = 21.18$, $p < .001$.

Intervention participants, relative to control participants, were also more likely to correctly answer the health-related math Problem 2 (i.e., COVID-19 fatality rates across time; Appendix Table E2) when controlling for gender, $b = 0.35$, $p = .007$, $OR = 1.41$, 95% CI [1.10, 1.82], or the larger set of covariates, $b = 0.40$, $p = .006$, $OR = 1.48$, 95% CI [1.12, 1.98]. Across conditions, more participants who incorrectly answered Problem 2 selected the WNB response than the other incorrect response, 77.66% versus 22.34%, $\chi^2(1) = 157.56$, $p < .001$.

Based on our preregistered analytic plan, we tested the interaction between magnitude knowledge (PAE) and condition, because people may benefit more from the educational intervention if they began the experiment with lower mathematical skills, thus had more "room to grow." Prior to testing interactions, we mean centered continuous variables. However, this interaction was not significant, did not improve model fit, and is not discussed further.

We also explored whether math attitudes and math anxiety moderated the effect of condition on the likelihood of correctly answering the health problems. In the simplified model with only gender as a covariate, math attitudes marginally moderated the effect of condition on the likelihood of correctly answering postintervention Problem 1. In our model with all covariates, math attitudes moderated the effect of condition for postintervention Problem 1 accuracy, but not Problem 2 accuracy. There was no interaction between condition and math anxiety on postintervention Problem 1 or Problem 2.

To further explore the math attitudes by condition interaction, we examined the effect of dummy coded condition (1 = experimental, 0 = control) on the likelihood of answering the problem correctly when math attitudes were more favorable (+1 *SD*) or less favorable (−1 *SD*). We found that the effect of training condition was larger when math attitudes were less favorable, $b = 1.72$, $SE = 0.20$, $p < .001$, $OR = 5.53$ [3.75, 8.27], compared to more favorable, $b = 1.06$, $SE = 0.22$, $p < .001$, $OR = 2.88$ [1.89, 4.44]. That is, those with less favorable math attitudes benefited more from training than those with more favorable attitudes.

## WNB in Strategy Reports Postintervention

In addition to accuracy, we examined open-ended strategy reports to assess whether the intervention improved not only participants' case fatality-rate calculation accuracy, but also decreased the likelihood that they reported a strategy consistent with WNB. We conducted logistic regressions predicting WNB strategy use on the postintervention health-related math problems.

As hypothesized, participants in the intervention condition were less likely to report WNB errors on postintervention Problem 1 (i.e., COVID-19 versus the flu) as compared to the control condition (observed frequencies: intervention = 10.08% versus control = 18.69%). See Appendix Table E3.

For health-related math Problem 2 (see Appendix Table E4), participants were less likely to report WNB errors in the intervention than in the control group (observed frequencies: intervention: 19.24% versus control: 23.36%). See Appendix Table E4. Across the intervention and control group, the proportion of people who reported WNB strategies was greater among those who answered the health-related math problems incorrectly compared to correctly: Problem 1, $\chi^2(1) = 187.78$, $p < .001$ (observed frequencies: 35.91% for incorrect versus 5.99% for correct) and Problem 2, $\chi^2(1) = 96.55$, $p < .001$ (observed frequencies: 27.97% for incorrect versus 2.36% for correct).

These strategy reports provide insight into *why* the intervention resulted in higher accuracy on the postintervention health-related math problems: fewer intervention participants explicitly mentioned WNB errors in their open-ended strategy reports than control participants. Prior to the intervention, participants in both conditions explicitly described WNB errors in their open-ended strategy reports with comparable frequency. This suggests that intervention participants obtained procedural and/or conceptual knowledge from the intervention that facilitated the use of more accurate strategies to solve the problems.

## Effect of the Intervention on Risk Perceptions, Worry, Affect, and Behavior

Next, we explored the immediate and long-term effects of the intervention on participants' COVID-19 risk perceptions, worry, positive and negative affect, and on reports of recommended health behaviors in daily diaries completed in the 10 days following the intervention. Interested readers can find the details of these models in the Supplement. It is important for future research to replicate these preliminary findings pertaining to the effect of the intervention on risk perceptions and worry.

[3] At baseline, controlling for gender, participants in both conditions were equally likely to mention whole number bias errors in their strategy reports across conditions, $b = 0.03$, $p = .805$, $OR = 1.03$, 95% CI [0.81, 1.32]. Notably, those who responded incorrectly were more likely to report a whole number bias error in their pretest strategy reports than those who answered correctly (47.6% vs. 2.5%), $\chi^2(1) = 317.09$, $p < .001$.

[4] The reason accuracy is this high in the control group is a question for future research.

**Table 5**
*Frequency of Incorrect and Correct Responses on the Health-Related Math Problems*

| Problem | Condition | # Incorrect | # Correct | $\chi^2$ |
|---|---|---|---|---|
| Postintervention Problem 1 (COVID-19 versus flu) | Control | 257 (40%) | 385 (60%) | 91.63, $p < .001$ |
| | Intervention | 105 (16%) | 550 (84%) | |
| Postintervention Problem 2 (COVID-19 fatality rates over time) | Control | 493 (77%) | 149 (23%) | 5.35, $p = .021$ |
| | Intervention | 465 (71%) | 190 (29%) | |

*Note.* COVID-19 = coronavirus disease. Chi-square tests are on observed frequencies only.



### *Immediate Effects Postintervention*

Neither risk perceptions nor worry differed across the intervention and control groups (Supplemental Tables 2 and 3). Negative affect, controlling for gender, did not differ by condition either (Supplemental Table 4). However, participants reported greater positive affect in the intervention relative to the control group, $b = 0.21$, $t(1287) = 2.82$, $p = .005$, indicating that the intervention did not cause immediate distress, but did *increase* positive affect (Supplemental Table 5). Therefore, it is unlikely that we scared participants into rating risks and worry higher after engaging in the educational intervention.

### *Effects of the Intervention Over Time*

Next, we conducted exploratory analyses examining the effects of the intervention on risk perceptions and worry across 10 days of the daily diary (Supplemental Figures 2 and 3 and Supplemental Tables 6 and 7). To summarize, when we plotted risk perceptions and worry by condition, we observed *nonlinear change*. Therefore, we explored the initial (days 1–3), middle (days 4–6), and final (days 7–10) periods of the diary assessment. Individuals in the intervention, relative to the control group, reported somewhat greater perceived risk during the middle period, $b = 0.39$, $SE = 0.20$, $p = .051$. However, perceived risk in the initial period and final period did not differ by condition. For worry, individuals in the intervention, relative to the control condition, reported greater worry in the initial period, $b = 0.43$, $SE = 0.18$, $p = .012$. For readers who are interested in the effect of the intervention on affect across 10 days, full details of our analyses can be found in the Supplement.

### *Engagement in Preventive Health Behaviors*

Finally, we explored the association between the intervention and recommended preventive health behaviors (per the U.S. CDC) reported over 10 days. There was no effect of condition on behaviors (Supplemental Table 10), but it is important to keep in mind that behavior change was not the main target of the intervention, whereas providing adults with procedural and conceptual knowledge of how to calculate case-fatality rates was the main goal.

Health behaviors are driven by numerous factors, including but not limited to, habits, emotion, socioeconomic status, health literacy, and health cognitions (Glanz & Bishop, 2010; Michie et al., 2011; Stokols, 1996). Importantly, some COVID-19 cognitions (e.g., susceptibility to the virus and perceptions of virus severity) did not directly predict preventive health behaviors in the control group (Coifman et al., 2021). However, it is an open question as to whether risk perceptions may have been more predictive of health behaviors at a later point in the pandemic. Future research can investigate whether a longer lasting intervention, or one that involves distributed practice (Dunlosky et al., 2013), might have a downstream effect on health behaviors.

## Math Anxiety Was Associated With Health-Related Problem-Solving Accuracy, Strategy Reports, Risk Perceptions, Worry, and Negative Affect at Baseline and During Daily Diaries

Our previous research has indicated that math anxiety is associated with objective and subjective numeracy (Choi et al., 2020). Over and above the effect of condition, math anxiety accounted for unique variance in logistic regressions on the outcome of health-related math problem-solving accuracy (Problem 2, $b = -0.08$) and WNB strategy reports (Problem 1, $b = 0.10$). See Appendix Tables E1–E4. Of note, math anxiety was also a significant positive predictor of risk perceptions ($b = 0.07$), worry ($b = 0.09$) and negative affect ($b = 0.23$) immediately postintervention during the baseline data collection session *and* across 10 days (risk perceptions, $b = 0.10$, worry, $b = 0.12$, and negative affect, $b = 0.09$; Supplemental Tables S1 and S8). These findings underscore the importance of considering participants' pre-existing math anxiety, and how to reduce it, as they reason about their perceptions of risk and worry pertaining to the global COVID-19 pandemic or any other health information expressed as rational numbers.

## Discussion

The COVID-19 pandemic remains an unprecedented health crisis. Throughout the last twenty months of the COVID-19 pandemic, statistics on COVID-19 infection, death, and now vaccination rates have updated by the minute, bombarding individuals with information that can be challenging to consume effectively. Our data clearly show that many adults have difficulty estimating the magnitude of ratios, specifically those that are relevant to interpreting COVID-19-related health risks. Our results align with the integrated theory of whole numbers and fractions development (Siegler et al., 2011), because we have shown that even adults fail to integrate their understanding of whole numbers and other ratios, such as fractions. We taught adults *how* to reason about ratios, a skill that could allow them to make more informed decisions about COVID-19 health statistics presented in the media and by health professionals. We also argue that this intervention may be applicable to other health information that is expressed as rational numbers. That is, we believe that this intervention is highly generalizable.

After our educational intervention, which relied on evidence-based approaches to improving rational number understanding, U.S.

adults were more accurate on health-related math problems comparing COVID-19 and flu case-fatality rates and also less likely to report WNB errors when reasoning about these case-fatality rates as compared to control participants. Strategy reports provided converging evidence of the effectiveness of the intervention. Those who engaged with the intervention were less likely to report WNB errors after they solved postintervention health-related math problems than those participants in the control group. Importantly, condition differences in accuracy and strategies held when covarying a range of sociodemographic and math factors. That is, even when we included variables previously shown to predict math performance, condition still significantly predicted health-related math accuracy and reported use of WNB strategies. Interestingly, we found that those participants who rated their math attitudes as less favorable benefited *more* from training than those with more favorable attitudes. Future work should investigate the nature of this interaction, as it suggests that in health contexts, individuals who negatively evaluate their math skills might benefit most from an intervention in which they are prompted to engage with mathematical computations rather than avoid them.

Clearly, much has changed about the public's understanding of COVID-19, since these data were collected at the beginning of the pandemic in March and April 2020. However, one thing that has *not* changed is that people can still commit WNB errors, as they attempt to comprehend health statistics, now with regard to vaccine rates across states and countries.

In summary, because there were no condition differences in health-related math problem-solving accuracy and WNB strategy reports at pretest (other than gender, and this individual difference variable was controlled for in subsequent models), we can conclude that the intervention *caused* post-test differences relating to accuracy in the health-related math problems. Specifically, individuals randomized to the intervention condition were more accurate on the health-related math problems and were less likely to report using WNB strategies.

## Limitations and Future Directions

Here, we describe limitations and ways to address these limitations in future lines of research. One limitation is that participants at the time of data collection could not know their specific, exact COVID-19 risk. Of course, any particular individual may be more or less at risk of death from COVID-19 than the population average. At the time of data collection, personalized risk estimates were not readily available, and media reports focused on the case-fatality rates of COVID-19 in general. Future work could help adults seek out and interpret more personalized risk information. However, more accurate knowledge about the infection rate, case-fatality rate, or vaccination rate of a given population (e.g., country, state, county, city, school district, etc.) could motivate individuals to engage in prevention behavior to protect the broader community.

Another limitation of the present study is that we did not assess calculator use and this may have impacted group differences. However, in similar studies (Mielicki et al., 2021; Scheibe et al., 2021), effects of mathematical problem-solving are the same when people who used calculators are excluded from analyses. In addition, when interpreting statistics in the media, many people presumably have access to calculators, which they may use (correctly or not) to better understand health statistics presented as rational numbers. If participants used a calculator in our study, they likely did so because they realized that computation requiring a calculator was necessary to *precisely* compare the flu case-fatality rate to the COVID-19 case-fatality rate, as opposed to considering numerators or denominators in isolation.

The fact that we compared our educational intervention against a business-as-usual control (i.e., no instruction) could be considered a limitation of the present study. However, this was our first attempt at "moving the dial" on rational number understanding during an emerging health crisis. Future research could include an "active" control condition in which participants are taught how to extract the appropriate numbers from a 2 × 2 contingency table or a word problem, plug them into a calculator by dividing the numerator by the denominator, and transform the resulting proportion (i.e., decimal) to a percentage by multiplying by 100. We would expect that teaching about calculation might facilitate problem-solving on very similar problems, but may not help learners understand the general, conceptual issue of how to identify an appropriate ratio from a 2 × 2 contingency table or a word problem. Our team is already beginning to examine some of these empirical questions (Mielicki et al., 2021).

A practical limitation of the present study was that it would have been ideal to include more than three health-related math problems postintervention to more extensively assess uptake of the intervention. However, completing word problems, such as the ones created for the present study, is time-intensive, and participants become frustrated after completing so many items, particularly since math items are often disliked and considered difficult and anxiety-provoking (Fagerlin et al., 2007; Sidney et al., 2021). In a follow-up study with undergraduate students, which was conducted in Fall 2020 (Mielicki et al., 2021), we created a larger battery of health-related math problems pertaining to COVID-19. Cronbach's α was approximately .6, suggesting that context differences among individual problems (e.g., topics or numbers involved in the problems) likely *do* matter for accuracy. Our results from that study—that is, relatively low reliability across multiple problems—provide support for our choice of analytic approach in the present study—analyzing one problem at a time with a logistic regression.

One potential route for future investigations is to determine whether individual differences in division skill predict performance on the health-related math problems, in which division is a critical operation. In a large sample (Siegler et al., 2012), both division performance and number-line estimation precision were important predictors of students' Algebra performance 5–6 years later, even when controlling for other important demographic factors. Therefore, it is possible that division ability *could* play a unique role in case-fatality rate calculation as well. Note, however, that PAE (|actual estimate|/scale of the number line) was the strongest predictor for the postintervention health-related problems when condition, math anxiety, math attitudes, and pretest health-related math problem solving were included simultaneously as predictors in the models for the present study. It is unclear whether division skills would also be a significant predictor when PAE, a strong proxy of overall math ability, already accounted for a high proportion of unique variance in problem-solving accuracy. Division and fractions are intricately linked, and fraction estimation PAE may already capture important variability in division skill.

## Additional Considerations

### *The Role of Gender in Health-Related Math Problem Solving*

An additional factor that was related to health-related math outcomes was gender: men were more likely than women to correctly answer the pretest health-related math problem ($OR$ = 1.98, 95% CI [1.59, 2.48]), possibly because men have been shown to have more precise magnitude understanding than women (Hutchison et al., 2019; Rivers et al., 2021; Thompson & Opfer, 2008). Future research should assess whether these gender differences in magnitude understanding in health-related math contexts are the product of differential spatial abilities (Newcombe et al., 2019; Rivers et al., 2021) and/or formal and informal learning experiences, such as early math and spatial talk in the home environment (Halpern et al., 2007; Levine et al., 2010; Pruden et al., 2011).

### *Interrelations Among Math Skills, Math Anxiety and Attitudes, and Health-Related Math Problem-Solving Accuracy*

The current interdisciplinary research team included experts in mathematical cognition, risk perceptions, and emotions. Each member of the research team viewed the COVID-19 pandemic through their own disciplinary lens to provide new insights on the relations between mathematical cognition, health cognition, and emotions in the context of the pandemic. We examined the impact of individual difference factors known to impact math performance in other lab-based math cognition (Sidney, Thalluri, et al., 2019) and health-related math studies (Choi et al., 2020; Mielicki et al., 2021; Woodbury et al., under review), on real-world COVID-19 risk perceptions and worry.

Not only did we find that our educational intervention impacted problem solving in the COVID-19 context, notably, math anxiety was also a unique, consistent, and strong predictor of risk perceptions, worry, and negative affect immediately after the intervention *and* across the 10-day follow-up. Therefore, to the extent that health statistics are communicated to the public in the form of symbolic (i.e., Arabic numerals) or nonsymbolic ratios (i.e., charts and graphs), math anxiety could play a role in predicting health-related math problem solving and health decision making. This finding highlights the ongoing need to understand how people's experiences with math contribute to their health and to find effective ways to minimize any impact of math anxiety on health numeracy in everyday scenarios.

## Conclusion and Practical Implications

In summary, our novel educational intervention involved training about COVID-19 case-fatality rates to diminish WNB—a pervasive mathematical misconception—to train adults to accurately interpret vital health information within the context of a global health crisis. This is the first study, to the best of our knowledge, in which WNB was shown to affect real-world health cognitions. Our study harnessed interdisciplinary, evidence-based practices from cognitive science to inform COVID-19 health communication. The intervention could be promising for any scenario in which rational-number information is involved and WNB errors are prevalent. Indeed, public health messaging is rife with potential for misinterpretation, from rates relating to disease risk (e.g., "1 in 8 women" will develop breast cancer) to vital health information relating to disease treatment (e.g., considering side-effect rates of drugs; Waters et al., 2007).

It might also be helpful for the media to present health statistics as easier-to-understand percentages accompanied by number line visuals that reinforce the magnitude of the health statistic, given the current findings and those reported by Mielicki et al. (2021). However, we want to underscore that more research is needed before we can make stronger prescriptive recommendations for policy change (see Robinson et al., 2013; Robinson & Levin, 2019, for arguments about going too far beyond one's data).

Based on our findings, it is important to help people understand the holistic magnitude of rates when rational numbers are included in health messaging. Interventions that facilitate improved interpretation and understanding of vital health statistics have the potential to make a positive impact on health decision making. As evidenced in our intervention, this was highly relevant during the current unprecedented COVID-19 public health crisis, but our intervention also has considerable potential for future health contexts, or other situations (i.e., considering interest rates on mortgages) in which people must engage with numerical information expressed as rational numbers.

## References

Alibali, M. W., & Sidney, P. G. (2015). Variability in the natural number bias: Who, when, how, and why. *Learning and Instruction*, *37*(5), 56–61. https://doi.org/10.1016/j.learninstruc.2015.01.003

Alonso-Diaz, S., Cantlon, J. F., & Piantadosi, S. T. (2019, July). *Intrinsic whole number bias in an indigenous population* [Poster presentation]. 41st Annual Conference of the Cognitive Science Society, Montreal, Quebec, Canada.

Ashcraft, M. H. (2002). Math anxiety: Personal, educational, and cognitive consequences. *Current Directions in Psychological Science*, *11*(5), 181–185. https://doi.org/10.1111/1467-8721.00196

Ashcraft, M. H., & Krause, J. A. (2007). Working memory, math performance, and math anxiety. *Psychonomic Bulletin & Review*, *14*(2), 243–248. https://doi.org/10.3758/BF03194059

Behrend, T. S., Sharek, D. J., Meade, A. W., & Wiebe, E. N. (2011). The viability of crowdsourcing for survey research. *Behavior Research Methods*, *43*(3), 800–813. https://doi.org/10.3758/s13428-011-0081-0

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(5), 1860–1863. https://doi.org/10.1073/pnas.0910967107

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology*, *54*(1), 579–616. https://doi.org/10.1146/annurev.psych.54.101601.145030

Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. Guilford Press.

Braithwaite, D. W., & Siegler, R. S. (2018). Developmental changes in the whole number bias. *Developmental Science*, *21*(2), Article e12541. https://doi.org/10.1111/desc.12541

Braithwaite, D. W., & Siegler, R. S. (2020). Putting fractions together. *Journal of Educational Psychology*. Advance online publication. https://doi.org/10.1037/edu0000477

Brewer, N. T., Chapman, G. B., Gibbons, F. X., Gerrard, M., McCaul, K. D., & Weinstein, N. D. (2007). Meta-analysis of the relationship between risk perception and health behavior: The example of vaccination. *Health Psychology*, *26*(2), 136–145. https://doi.org/10.1037/0278-6133.26.2.136

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*(1), 116–131. https://doi.org/10.1037/0022-3514.42.1.116

Centers for Disease Control and Prevention. (2020, April 17). *2019–2020 U.S. flu season: Preliminary burden estimates*. https://www.cdc.gov/flu/about/burden/preliminary-in-season-estimates.htm

Chapman, G. B., & Coups, E. J. (2006). Emotions and preventive health behavior: Worry, regret, and influenza vaccination. *Health Psychology*, *25*(1), 82–90. https://doi.org/10.1037/0278-6133.25.1.82

Chew, L. D., Griffin, J. M., Partin, M. R., Noorbaloochi, S., Grill, J. P., Snyder, A., Bradley, K. A., Nugent, S. M., Baines, A. D., & Vanryn, M. (2008). Validation of screening questions for limited health literacy in a large VA outpatient population. *Journal of General Internal Medicine*, *23*(5), 561–566. https://doi.org/10.1007/s11606-008-0520-5

Choi, S. S., Taber, J. M., Thompson, C. A., & Sidney, P. G. (2020). Math anxiety, but not induced stress, is associated with objective numeracy. *Journal of Experimental Psychology: Applied*, *26*(4), 604–619. https://doi.org/10.1037/xap0000268

Coelho, G. L. H., Hanel, P. H. P., & Wolf, L. K. (2018). The very efficient assessment of need for cognition: Developing a six-item version. *Assessment*, *27*(8), 1–16. https://doi.org/10.1177/1073191118793208

Coifman, K. G., Disabato, D. J., Aurora, P., Seah, T. H. S., Simonovic, N., Foust, J. L., Sidney, P. G., Thompson, C. A., & Taber, J. M. (2021). What drives preventive health behavior during a global pandemic? Emotion and worry. *Annals of Behavioral Medicine*, *55*(8), 791–804. https://doi.org/10.1093/abm/kaab048

Coifman, K. G., Flynn, J. J., & Pinto, L. A. (2016). When context matters: Negative emotions predict psychological health and adjustment. *Motivation and Emotion*, *40*(4), 602–624. https://doi.org/10.1007/s11031-016-9553-y

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin numeracy test. *Judgment and Decision Making*, *7*(1), 25–47.

Cranford, J. A., Shrout, P. E., Iida, M., Rafaeli, E., Yip, T., & Bolger, N. (2006). A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies detect change reliably? *Personality and Social Psychology Bulletin*, *32*(7), 917–929. https://doi.org/10.1177/0146167206287721

DeWolf, M., & Vosniadou, S. (2015). The representation of fraction magnitudes and the whole number bias reconsidered. *Learning and Instruction*, *37*, 39–49. https://doi.org/10.1016/j.learninstruc.2014.07.002

Di Donato, V., McKenzi, S., & Borghese, L. (2020). *Italy's coronavirus death toll passes 10,000. Many are asking why the fatality rate is so high*. https://www.cnn.com/2020/03/28/europe/italy-coronavirus-cases-surpass-china-intl/index.html

Dowker, A., Sarkar, A., & Looi, C. Y. (2016). Mathematics Anxiety: What have we learned in 60 years? *Frontiers in Psychology*, *7*, Article 508. https://doi.org/10.3389/fpsyg.2016.00508

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring numeracy without a math test: Development of the Subjective Numeracy Scale. *Medical Decision Making*, *27*(5), 672–680. https://doi.org/10.1177/0272989X07304449

Faust, J., & Del Rio, C. (2020b). Relative disease burdens of COVID-19 and seasonal influenza in New York City, February 1–April 18, 2020. *MedRxiv*. https://doi.org/10.1101/2020.04.22.20073551

Faust, J. M. (2020, April 28). Comparing COVID-19 deaths to Flu deaths is like comparing apples to oranges. *Scientific American's Newsletters*. https://blogs.scientificamerican.com/observations/comparing-covid-19-deaths-to-flu-deaths-islike-comparing-apples-to-oranges

Faust, J. S., & Del Rio, C. (2020a). Assessment of deaths from COVID-19 and from seasonal influenza. *JAMA Internal Medicine*, *180*(8), 1045–1046. https://doi.org/10.1001/jamainternmed.2020.2306

Fazio, L. K., Bailey, D. H., Thompson, C. A., & Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *Journal of Experimental Child Psychology*, *123*, 53–72. https://doi.org/10.1016/j.jecp.2014.01.013

Fazio, L. K., DeWolf, M., & Siegler, R. S. (2016). Strategy use and strategy choice in fraction magnitude comparison. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(1), 1–16. https://doi.org/10.1037/xlm0000153

Ferrer, R. A., Klein, W. M., Persoskie, A., Avishai-Yitshak, A., & Sheeran, P. (2016). The tripartite model of risk perception (TRIRISK): Distinguishing deliberative, affective, and experiential components of perceived risk. *Annals of Behavioral Medicine*, *50*(5), 653–663. https://doi.org/10.1007/s12160-016-9790-z

Fitzsimmons, C., Morehead, K., Thompson, C. A., Buerke, M., & Dunlosky, J. (2021). Can feedback, correct, and incorrect worked examples improve numerical magnitude estimation precision? *Journal of Experimental Education*. Advance online publication. https://doi.org/10.1080/00220973.2021.1891009

Fitzsimmons, C. J., Thompson, C. A., & Sidney, P. G. (2020). Do adults treat equivalent fractions equally? Adults' strategies and errors during fraction reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(11), 2049–2074. https://doi.org/10.1037/xlm0000839

Fuchs, L. S., Sterba, S. K., Fuchs, D., & Malone, A. S. (2016). Does evidence-based fractions intervention address the needs of very low-performing students? *Journal of Research on Educational Effectiveness*, *9*(4), 662–677. https://doi.org/10.1080/19345747.2015.1123336

Gilman, T. L., Shaheen, R., Nylocks, K. M., Halachoff, D., Chapman, J., Flynn, J. J., Matt, L. M., & Coifman, K. G. (2017). A film set for the elicitation of emotion in research: A comprehensive catalog derived from four decades of investigation. *Behavior Research Methods*, *49*(6), 2061–2082. https://doi.org/10.3758/s13428-016-0842-x

Glanz, K., & Bishop, D. B. (2010). The role of behavioral science theory in development and implementation of public health interventions. *Annual Review of Public Health*, *31*(1), 399–418. https://doi.org/10.1146/annurev.publhealth.012809.103604

Gómez, D. M., Jiménez, A., Bobadilla, R., Reyes, C., & Dartnell, P. (2015). The effect of inhibitory control on general mathematics achievement and fraction comparison in middle school children. *ZDM*, *47*(5), 801–811. https://doi.org/10.1007/s11858-015-0685-4

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest*, *8*(1), 1–51. https://doi.org/10.1111/j.1529-1006.2007.00032.x

Hamdan, N., & Gunderson, E. A. (2017). The number line is a critical spatial-numerical representation: Evidence from a fraction intervention. *Developmental Psychology*, *53*(3), 587–596. https://doi.org/10.1037/dev0000252

Hauser, D. J., & Schwarz, N. (2016). Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods*, *48*(1), 400–407. https://doi.org/10.3758/s13428-015-0578-z

Hay, J. L., McCaul, K. D., & Magnan, R. E. (2006). Does worry about breast cancer predict screening behaviors? A meta-analysis of the prospective evidence. *Preventive Medicine*, *42*(6), 401–408. https://doi.org/10.1016/j.ypmed.2006.03.002

Horowitz, J., Bubola, E., & Povoledo, E. (2020). *Italy, pandemic's new epicenter, has lessons for the world*. https://www.nytimes.com/2020/03/21/world/europe/italy-coronavirus-center-lessons.html

Huang, P. (2020). *How the novel coronavirus and the flu are alike . . . and different*. https://www.npr.org/sections/goatsandsoda/2020/03/20/815408287/how-the-novel-coronavirus-and-the-flu-are-alike-and-different

Hutchison, J. E., Lyons, I. M., & Ansari, D. (2019). More similar than different: Gender differences in children's basic numerical skills are the exception not the rule. *Child Development*, *90*(1), e66–e79. https://doi.org/10.1111/cdev.13044

Johns Hopkins University and Medicine. (2020, May 27). *COVID-19 United States cases by county*. https://coronavirus.jhu.edu/us-map

Klein, W. M. P., & Ferrer, R. A. (2018). On being more amenable to threatening risk messages concerning close others (vis-a-vis the self). *Personality and Social Psychology Bulletin*, *44*(10), 1411–1423. https://doi.org/10.1177/0146167218769064

Levine, S. C., Suriyakham, L. W., Rowe, M. L., Huttenlocher, J., & Gunderson, E. A. (2010). What counts in the development of young children's number knowledge? *Developmental Psychology*, *46*(5), 1309–1319. https://doi.org/10.1037/a0019671

Lipkus, I. M., & Peters, E. (2009). Understanding the role of numeracy in health: Proposed theoretical framework and practical insights. *Health Education & Behavior*, *36*(6), 1065–1081. https://doi.org/10.1177/1090198109341533

McGinn, K. M., Lange, K. E., & Booth, J. L. (2015). A worked example for creating worked examples. *Mathematics Teaching in the Middle School*, *21*(1), 26–33. https://doi.org/10.5951/mathteacmiddscho.21.1.0026

Meyerson, P., & Tryon, W. W. (2003). Validating internet research: A test of the psychometric equivalence of internet and in-person samples. *Behavior Research Methods, Instruments, & Computers*, *35*(4), 614–620. https://doi.org/10.3758/BF03195541

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science: IS*, *6*(1), Article 42. https://doi.org/10.1186/1748-5908-6-42

Mielicki, M., Fitzsimmons, C., Schiller, L., Scheibe, D., Taber, J. M., Sidney, P., Matthews, P. G., Waters, E. A., Coifman, K., & Thompson, C. A. (2021, June 10). Adults' health-related problem solving is facilitated by number lines, but not risk ladders and icon arrays. *PsyArXiv*. https://doi.org/10.31234/osf.io/h3stw

Mills, M., Rahal, C., & Akimova, E. (2020). *Face masks and coverings for the general public: Behavioural knowledge, effectiveness of cloth coverings and public messaging*. The Royal Society and The British Academy. https://royalsociety.org/-/media/policy/projects/set-c/set-c-facemasks.pdf

Moss, J., & Case, R. (1999). Developing children's understanding of the rational numbers: A new model and an experimental curriculum. *Journal for Research in Mathematics Education*, *30*(2), 122–147. https://doi.org/10.2307/749607

Nelson, W., Reyna, V. F., Fagerlin, A., Lipkus, I., & Peters, E. (2008). Clinical implications of numeracy: Theory and practice. *Annals of Behavioral Medicine*, *35*(3), 261–274. https://doi.org/10.1007/s12160-008-9037-8

Netburn, D. (2020, March 5). The flu has killed far more people than coronavirus. So why all the frenzy about COVID-19? *Los Angeles Times*. https://www.latimes.com/science/story/2020-03-05/flu-killed-more-people-coronavirus-covid-19

Newcombe, N. S., Booth, J. L., & Gunderson, E. A. (2019). Spatial skills, reasoning, and mathematics. In J. Dunlosky & K. A. Rawson (Eds.), *Cambridge handbook of cognition and education* (pp. 100–123). Cambridge University Press. https://doi.org/10.1017/9781108235631.006

Ni, Y., & Zhou, Y. D. (2005). Teaching and learning fraction and rational numbers: The origins and implications of whole number bias. *Educational Psychologist*, *40*(1), 27–52. https://doi.org/10.1207/s15326985ep4001_3

Obersteiner, A., Van Dooren, W., Van Hoof, J., & Verschaffel, L. (2013). The natural number bias and magnitude representation in fraction comparison by expert mathematicians. *Learning and Instruction*, *28*(8), 64–72. https://doi.org/10.1016/j.learninstruc.2013.05.003

Opfer, J. E., & Devries, J. M. (2008). Representational change and magnitude estimation: Why young children can make more accurate salary comparisons than adults. *Cognition*, *108*(3), 843–849. https://doi.org/10.1016/j.cognition.2008.05.003

Opfer, J. E., & Siegler, R. S. (2007). Representational change and children's numerical estimation. *Cognitive Psychology*, *55*(3), 169–195. https://doi.org/10.1016/j.cogpsych.2006.09.002

Opfer, J. E., & Thompson, C. A. (2008). The trouble with transfer: Insights from microgenetic changes in the representation of numerical magnitude. *Child Development*, *79*(3), 788–804. https://doi.org/10.1111/j.1467-8624.2008.01158.x

Opfer, J. E., Thompson, C. A., & Kim, D. (2016). Free versus anchored numerical estimation: A unified approach. *Cognition*, *149*, 11–17. https://doi.org/10.1016/j.cognition.2015.11.015

Perez-Pena, R. (2020). *Virus hits Europe harder than China. Is that the price of an open society?* https://www.nytimes.com/2020/03/19/world/europe/europe-china-coronavirus.html

Peters, E. (2012). Beyond comprehension: The role of numeracy in judgments and decisions. *Current Directions in Psychological Science*, *21*(1), 31–35. https://doi.org/10.1177/0963721411429960

Portnoy, D. B., Kaufman, A. R., Klein, W. M., Doyle, T. A., & de Groot, M. (2014). Cognitive and affective perceptions of vulnerability as predictors of exercise intentions among people with type 2 diabetes. *Journal of Risk Research*, *17*(2), 177–193. https://doi.org/10.1080/13669877.2013.794153

Pruden, S. M., Levine, S. C., & Huttenlocher, J. (2011). Children's spatial thinking: Does talk about the spatial world matter? *Developmental Science*, *14*(6), 1417–1430. https://doi.org/10.1111/j.1467-7687.2011.01088.x

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Renkl, A. (2014). Toward an instructionally oriented theory of example-based learning. *Cognitive Science*, *38*(1), 1–37. https://doi.org/10.1111/cogs.12086

Rettner, R. (2020, May 14). *How does the new coronavirus compare with the flu? Live Science*. https://www.livescience.com/new-coronavirus-compare-with-flu.html

Reyna, V. F., & Brainerd, C. J. (2008). Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learning and Individual Differences*, *18*(1), 89–107. https://doi.org/10.1016/j.lindif.2007.03.011

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin*, *135*(6), 943–973. https://doi.org/10.1037/a0017327

Rittle-Johnson, B. (2017). Developing mathematics knowledge. *Child Development Perspectives*, *11*(3), 184–190. https://doi.org/10.1111/cdep.12229

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, *93*(2), 346–362. https://doi.org/10.1037/0022-0663.93.2.346

Rivers, M. L., Fitzsimmons, C. J., Fisk, S. R., Dunlosky, J., & Thompson, C. A. (2021). Gender differences in confidence during number-line estimation. *Metacognition and Learning*, *16*(1), 157–178. https://doi.org/10.1007/s11409-020-09243-7

Robinson, D. H., & Levin, J. R. (2019). Quackery in educational research. In J. Dunlosky & K. A. Rawson (Eds.), *The Cambridge handbook of cognition and education* (pp. 35–48). Cambridge University Press. https://doi.org/10.1017/9781108235631.003

Robinson, D. H., Levin, J. R., Schraw, G., Patall, E. A., & Hunt, E. B. (2013). On going(way) beyond one's data: A proposal to restrict recommendations for practice in primary educational research journals. *Educational Psychology Review*, *25*(2), 291–302. https://doi.org/10.1007/s10648-013-9223-5

Scheibe, D. A., Fitzsimmons, C. J., & Thompson, C. A. (2021). *Number lines assist students as they calculate final course grades* [Manuscript in preparation].

Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology*, *103*(4), 759–775. https://doi.org/10.1037/a0025140

Sheeran, P., Harris, P. R., & Epton, T. (2014). Does heightening risk appraisals change people's intentions and behavior? A meta-analysis of experimental studies. *Psychological Bulletin*, *140*(2), 511–543. https://doi.org/10.1037/a0033065

Sidney, P., & Thompson, C. A. (2019). Supporting children's analogical learning by "warming up" with relevant prior knowledge. *Current Directions in Psychological Science*, *28*(6), 619–625. https://doi.org/10.1177/0963721419870801

Sidney, P., Thompson, C. A., & Rivera, F. (2019). Number lines, but not area models, support children's accuracy and conceptual models of fraction division. *Contemporary Educational Psychology*, *58*(3), 288–298. https://doi.org/10.1016/j.cedpsych.2019.03.011

Sidney, P. G., & Alibali, M. W. (2017). Creating a context for learning: Activating children's whole number knowledge prepares them to understand fraction division. *Journal of Numerical Cognition*, *3*(1), 31–57. https://doi.org/10.5964/jnc.v3i1.71

Sidney, P. G., Thalluri, R., Buerke, M. L., & Thompson, C. A. (2019). Who uses more strategies? Linking mathematics anxiety to adults' strategy variability and performance on fraction magnitude tasks. *Thinking & Reasoning*, *25*(1), 94–131. https://doi.org/10.1080/13546783.2018.1475303

Sidney, P. G., Thompson, C. A., Fitzsimmons, C., & Taber, J. M. (2021). Children's and adults' math attitudes are differentiated by number type. *Journal of Experimental Education*, *89*(1), 1–32. https://doi.org/10.1080/00220973.2019.1653815

Siegler, R. S. (2016). Magnitude knowledge: The common core of numerical development. *Developmental Science*, *19*(3), 341–361. https://doi.org/10.1111/desc.12395

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, *75*(2), 428–444. https://doi.org/10.1111/j.1467-8624.2004.00684.x

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M. I., & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, *23*(7), 691–697. https://doi.org/10.1177/0956797612440101

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, *14*(3), 237–250. https://doi.org/10.1111/1467-9280.02438

Siegler, R. S., & Thompson, C. A. (2014). Numerical landmarks are useful—Except when they're not. *Journal of Experimental Child Psychology*, *120*(3), 39–58. https://doi.org/10.1016/j.jecp.2013.11.014

Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology*, *62*(4), 273–296. https://doi.org/10.1016/j.cogpsych.2011.03.001

Stokols, D. (1996). Translating social ecological theory into guidelines for community health promotion. *American Journal of Health Promotion*, *10*(4), 282–298. https://doi.org/10.4278/0890-1171-10.4.282

Taber, J. M., Aspinwall, L. G., Drummond, D. M., Stump, T. K., Kohlmann, W., Champine, M., Cassidy, P., & Leachman, S. A. (2021). Priority of risk (but not perceived magnitude of risk) predicts improved sun-protection behavior following genetic counseling for familial melanoma. *Annals of Behavioral Medicine*, *55*(1), 24–40. https://doi.org/10.1093/abm/kaaa028

Taber, J. M., Klein, W. M., Suls, J. M., & Ferrer, R. A. (2017). Lay awareness of the relationship between age and cancer risk. *Annals of Behavioral Medicine*, *51*(2), 214–225. https://doi.org/10.1007/s12160-016-9845-1

Taber, J. M., McQueen, A., Simonovic, N., & Waters, E. A. (2019). Adapting a self-affirmation intervention for use in a mobile application for smokers. *Journal of Behavioral Medicine*, *42*(6), 1050–1061. https://doi.org/10.1007/s10865-019-00028-1

Thompson, C. A., Mielicki, M., Rivera, F., Fitzsimmons, C. J., Scheibe, D. A., Sidney, P. G., Taber, J. M., & Waters, E. A. (2021). Leveraging math cognition to combat health innumeracy. *Perspectives on Psychological Science*. accepted for publication.

Thompson, C. A., & Opfer, J. E. (2008). Costs and benefits of representational change: Effects of context on age and sex differences in symbolic magnitude estimation. *Journal of Experimental Child Psychology*, *101*(1), 20–51. https://doi.org/10.1016/j.jecp.2008.02.003

Thompson, C. A., & Opfer, J. E. (2010). How 15 hundred is like 15 cherries: Effect of progressive alignment on representational changes in numerical cognition. *Child Development*, *81*(6), 1768–1786. https://doi.org/10.1111/j.1467-8624.2010.01509.x

Thompson, C. A., & Opfer, J. E. (2016). Learning linear spatial-numeric associations improves accuracy of memory for numbers. *Frontiers in Psychology*, *7*, Article 24. https://doi.org/10.3389/fpsyg.2016.00024

Thompson, C. A., Taber, J. M., Coifman, K. G., & Sidney, P. (2020, April 8). Math misconceptions may lead people to underestimate the true threat of COVID-19. *The Conversation*. https://theconversation.com/math-misconceptions-may-lead-people-to-underestimate-the-true-threat-of-covid-19-134520

Thompson, C. A., Taber, J. M., Fitzsimmons, C., & Sidney, P. (in press). Strategy reports involving attention to math are associated with accurate responses on a numeric health-related math problem. *Journal of Numerical Cognition*.

United States Census Bureau. (2018). *Quick facts*. https://www.census.gov/quickfacts/fact/table/US/SBO030212

Vamvakoussi, X. (2015). The development of rational number knowledge: Old topic, new insights. *Learning and Instruction*, *37*(6), 50–55. https://doi.org/10.1016/j.learninstruc.2015.01.002

Vamvakoussi, X., & Vosniadou, S. (2010). How many decimals are there between two fractions? Aspects of secondary school students' understanding of rational numbers and their notation. *Cognition and Instruction*, *28*(2), 181–209. https://doi.org/10.1080/07370001003676603

Van Hoof, J., Verschaffel, L., De Neys, W., & Van Dooren, W. (2020). Intuitive errors in learners' fraction understanding: A dual-process perspective on the natural number bias. *Memory & Cognition*, *48*(7), 1171–1180. https://doi.org/10.3758/s13421-020-01045-1

Van Hoof, J., Verschaffel, L., & Van Dooren, W. (2015). Inappropriately applying natural number properties in rational number tasks: Characterizing the development of the natural number bias through primary and secondary education. *Educational Studies in Mathematics*, *90*(1), 39–56. https://www.jstor.org/stable/43590259

Vosniadou, S. (2014). Examining cognitive development from a conceptual change point of view: The framework theory approach. *European Journal of Developmental Psychology*, *11*(6), 645–661. https://doi.org/10.1080/17405629.2014.921153

Walker, M. (2020, April 27). COVID-19 no worse than the flu? Hardly: New York City data show dramatic imbalance. *MedPage Today*. https://www.medpagetoday.com/infectiousdisease/covid19/86176

Waters, E. A., Weinstein, N. D., Colditz, G. A., & Emmons, K. M. (2007). Reducing aversion to side effects in preventive medical treatment decisions. *Journal of Experimental Psychology: Applied*, *13*(1), 11–21. https://doi.org/10.1037/1076-898X.13.1.11

Weinstein, N. D., Kwitel, A., McCaul, K. D., Magnan, R. E., Gerrard, M., & Gibbons, F. X. (2007). Risk perceptions: Assessment and relationship to influenza vaccination. *Health Psychology*, *26*(2), 146–151. https://doi.org/10.1037/0278-6133.26.2.146

Weller, J. A., Dieckmann, N. F., Tusler, M., Mertz, C. K., Burns, W. J., & Peters, E. (2013). Development and testing of an abbreviated numeracy scale: A Rasch analysis approach. *Journal of Behavioral Decision Making*, *26*(2), 198–212. https://doi.org/10.1002/bdm.1751

Woodbury, L. H., Fitzsimmons, C. J., Taber, J. M., Schiller, L. K., Mielicki, M. K., & Thompson, C. A. (under review). *Impact of presentation format on health-risk estimates* [Manuscript under review].

World Health Organization. (2020). *Novel coronavirus (2019-nCoV): Situation report—13*. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200202-sitrep-13-ncov-v3.pdf?sfvrsn=195f4010_6

Yan, H. (2020, May 1). Critics said the flu kills more than coronavirus. Why that's not a fair comparison—and now, it's not even true. *CNN*. https://www.cnn.com/2020/05/01/health/flu-vs-coronavirus-deaths/index.html

Yang, S., Cao, P., Du, P., Wu, Z., Zhuang, Z., Yang, L., Yu, X., Zhou, Q., Feng, X., Wang, X., Weiguo, L., Liu, E., Chen, J., Chen, Y., & He, D. (2020). Early estimation of the case fatality rate of COVID-19 in mainland China: A data-driven analysis. *Annals of Translational Medicine*, *8*(4), Article 128. https://doi.org/10.21037/atm.2020.02.66

Yu, S., Kim, D., Mielicki, M. K., Fitzsimmons, C., Thompson, C. A., & Opfer, J. E. (2020). From integers to fractions: Developing a coherent understanding of proportional magnitude. In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of the 41st annual conference of the cognitive science society*. Cognitive Science Society.

# Appendix A

## Exclusion Criteria

Due to the nature of online sampling, we excluded participants who showed random or inattentive responding during the baseline survey and the daily diary. Of the 2,693 people who consented to participate, 832 were excluded for completing the survey in less than 13.33 min, or ⅓ of the length of time (40 min) that it took research assistants to complete the survey on average during piloting. Of the remaining 1,861 participants, 305 were excluded for failing at least one of two attention checks (i.e., Select answer B; Gilman et al., 2017), and two were excluded for having completed the survey more than once. Thus, our preanalytic sample included 1,554 people.

Additional preregistered exclusion criteria were aimed at identifying "poor responders" on the baseline math tasks. For each number line estimation task (i.e., fractions, whole number frequencies, and percentages), a participant was flagged for exclusion if they had number-line estimation precision (PAE, see below) that had a standard deviation of .1 or less, or if 80% of their responses within a particular numerical range were above 95% or below 5% of the line (Fitzsimmons et al., 2021; Sidney et al., 2021). That is, these participants had anchored their estimates at the endpoints of the number line and only placed their estimates within 10% of the number line on 80% or more of the trials. In the equivalence task, participants were flagged as poor responders for responding true or false to all items. Twenty participants were flagged as poor responders on two or more of these math tasks and excluded from analyses. Of the remaining 1,534 participants, we excluded eight participants who provided nonsensical answers to at least one of two open-ended items at the end of the study (e.g., "Is there anything else you would like to add regarding your thoughts about COVID-19?). We confirmed that these eight participants also provided nonsensical responses in their open-ended health decision-making strategy reports, suggesting that they were not engaging with any of the open-ended items. Of these 1,526 participants, the 1,297 who answered all four forced-choice health decision-making questions comprised our analytic sample for the primary analyses.

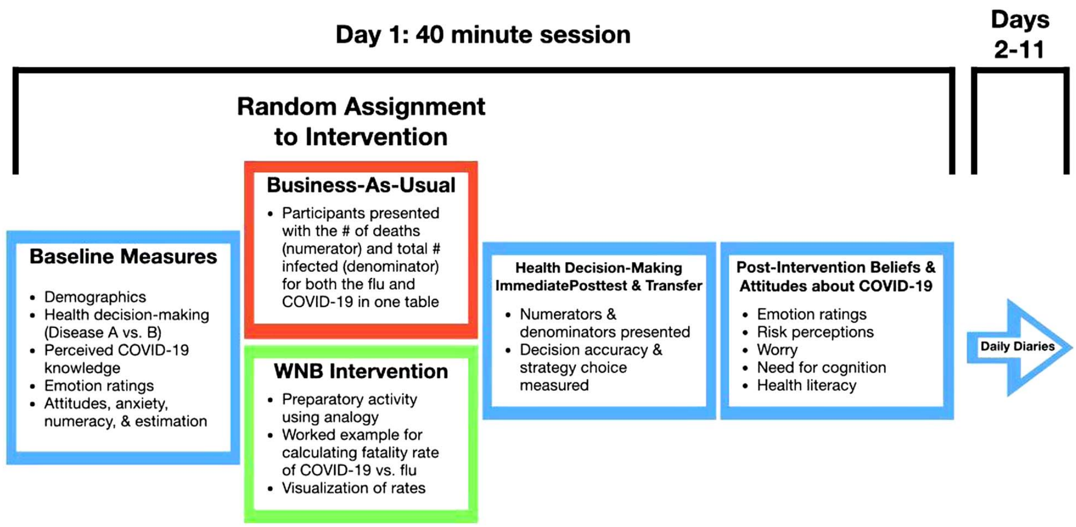
Our analytic sample differed from those who consented to participate, yet were excluded, in terms of their age, $t(2680) = 10.15$, $p < .001$, gender, $\chi^2(1) = 6.62$, $p = .01$, employment status, $\chi^2(6) = 50.32$, $p < .001$, race/ethnicity (coded as: white versus any other race), $\chi^2(1) = 182.47$, $p < .0001$, reported income, $\chi^2(9) = 71.06$, $p < .001$, number of math courses taken, $t(2633.5) = 11.70$, $p < .001$, accuracy on an objective numeracy question (Berlin numeracy scale), $\chi^2(1) = 95.44$, $p < .0001$, and accuracy on the baseline health decision-making problem, $\chi^2(1) = 45.75$, $p < .0001$. That is, those who were excluded were more likely to be younger, female, students, or self-employed, but less likely to be retired or employed for wages, be nonwhite, report lower income, have taken fewer math courses, and be incorrect on the objective numeracy and baseline health decision-making question. However, because our exclusion criteria aimed to identify nonhuman responding, these comparisons should be interpreted with caution.

Of the 1,297 participants from the baseline analytic sample, 709 participants completed the daily diaries over the next 10 days. 64 participants were excluded from the daily-diary sample, because they responded to signals less than 2 *SD* below the sample mean ($M = 7.58$, $SD = 3.11$), per convention (Bolger et al., 2003). Our analytic sample for the daily diary analyses was 627, because an additional 18 individuals failed embedded accuracy checks commonly used in online research (Gilman et al., 2017). The overall rate of compliance was acceptable, attrition was as expected and consistent with prior research, and there were approximately 4,703 diary signals available for analysis (Bolger & Laurenceau, 2013).

(*Appendices continue*)

## Appendix B

### Survey Flow




*Note*. See the online article for the color version of this figure.

## Appendix C

### Health Decision-Making Strategy Usage

Strategy report coding scheme with definitions, examples, and proportion of participants in each condition that used each strategy for each health decision-making item

| Code | Definition | Examples | Proportion of control participants who used strategy | Proportion of intervention participants who used strategy |
|---|---|---|---|---|
| Larger numbers | Indicated using larger numbers to guide decision but doesn't specify which numbers (numerator or denominator). | • Because the number is bigger<br>• I just compared the numbers and picked the highest one<br>• Look at the biggest number | B: .04<br>P1: .03<br>P2: .10<br>P3: .04 | B: .04<br>P1: .02<br>P2: .08<br>P3: .04 |
| Numerators | Mentioned the numerators (e.g., number of deaths) in isolation when comparing diseases. | • Most number dead<br>• The number of deaths from Disease B is significantly higher than Disease A.<br>• How many died from it | B: .23<br>P1: .16<br>P2: .11<br>P3: .07 | B: .24<br>P1 .07<br>P2: .10<br>P3: .07 |
| Denominators | Mentioned the denominators (e.g., number of people infected) in isolation when comparing diseases. | • Disease B has more people infected<br>• More cases and so on<br>• The population is lower in Italy | B: .14<br>P1: .08<br>P2: .08<br>P3: .13 | B: .15<br>P1: .03<br>P2: .06<br>P3: .13 |
| Rate | Indicated calculating a rate or comparing numbers, including transformations. | • I divided the cases into the deaths<br>• They are close to the same percentage | B: .48<br>P1: .47<br>P2: .38<br>P3: .41 | B: .50<br>P1: .45<br>P2: .42<br>P3: .43 |

(*Appendices continue*)

**Appendix C** (*continued*)

| Code | Definition | Examples | Proportion of control participants who used strategy | Proportion of intervention participants who used strategy |
|---|---|---|---|---|
| | | • The second is one per thousand infected | | |
| Math error | Indicated calculating a rate (e.g., by dividing the number of deaths by the number infected; dividing the number of infected by the total population), but made some type of math error (e.g., incorrect rounding). | • I tried to divide the number of people who had the disease by the number of people who died<br>• .056 (instead of .0056 for far transfer)<br>• Disease B is more fatal, because 10% of the total # of infected individuals died (instead of 1% for baseline) | B: .04<br>P1: .03<br>P2: .02<br>P3: .02 | B: .07<br>P1: .01<br>P2: .03<br>P3: .03 |
| Ambiguous math | Mentioned math terminology or some type of calculation. If the participant indicated a rate but was not specific about how it was calculated, this code was used together with the "rate" code above. | • According to the statistics<br>• I did calculations on both countries<br>• Using math | B: .34<br>P1: .33<br>P2: .37<br>P3: .30 | B: .32<br>P1: .39<br>P2: .39<br>P3: .33 |
| Condition-specific information | Used information given in the intervention condition. | • Rotting apples compared to trees.<br>• You stated COVID-19 was 41 times more fatal<br>• It says on the graph | B: .00<br>P1: .00<br>P2: .00<br>P3: .00 | B: .00<br>P1: .15<br>P2: .02<br>P3: .00 |
| News or media | Mentioned getting information from the news/media or a source other than the information provided in this study | • What I heard on the news reports rather than the above numbers<br>• United Nations information<br>• According to CNN, China has got in control of the spread of the Coronavirus. | B: .00<br>P1: .02<br>P2: .03<br>P3: .06 | B: .00<br>P1: .02<br>P2: .03<br>P3: .06 |
| Personal beliefs or opinions | Relied on personal beliefs or opinions, including facts (or myths) about COVID-19 not included in the prompt. | • Because it is targeting older people with health issues and with not the best health care systems.<br>• Any disease that can kill a person is fatal<br>• It started in China and has been the main location ever, since the disease started | B: .09<br>P1: .13<br>P2: .05<br>P3: .06 | B: .07<br>P1: .10<br>P2: .05<br>P3: .07 |
| Other | Mentioned guessing; provided a response that gave no insight into reasoning; mentioned needing more information to decide. | • Process of elimination<br>• Just do<br>• I can't determine which is the deadliest disease with just the information given | B: .07<br>P1: .10<br>P2: .13<br>P3: .13 | B: .06<br>P1: .10<br>P2: .12<br>P3: .12 |
| Nonsense | Responded with strings of letters or words that were nonsensical | • Ddufiffm<br>• How to see real good feel-ing in there how volume in how to earl | B: .01<br>P1: .01<br>P2: .02<br>P3: .02 | B: .01<br>P1: .02<br>P2: .01<br>P3: .02 |
| No response | Did not provide a response | | B: .02<br>P1: .02<br>P2: .03<br>P3 .03 | B: .02<br>P1: .02<br>P2: .03<br>P3: .03 |

*Note*. A response could receive more than one type of code. The first three codes (larger numbers, numerators, denominators) are considered to be consistent with whole number bias (WNB). B = baseline; P1 = postintervention Problem 1; P2 = postintervention Problem 2; P3 = postintervention Problem 3; COVID-19 = coronavirus disease.

(*Appendices continue*)

# Appendix D

## Daily Diary Behaviors

Please indicate whether you *performed any of the following actions or behaviors* in the past day:

| Behavior | No | Yes | Does not apply |
|---|---|---|---|
| Washed hands regularly or used antibacterial products | ○ | ○ | ○ |
| Purchased antibacterial or cleaning products | ○ | ○ | ○ |
| Looked for information about health or medical topics from any source | ○ | ○ | ○ |
| Looked for information about COVID-19 from any source | ○ | ○ | ○ |
| Shared information online or in-person with others about COVID-19 | ○ | ○ | ○ |
| Spoke to or contacted a medical professional about your health | ○ | ○ | ○ |
| Purchased extra food for home | ○ | ○ | ○ |
| Used a surgical mask to cover your face | ○ | ○ | ○ |
| Avoided public transportation | ○ | ○ | ○ |
| Avoided contact with people other than those who live in your household | ○ | ○ | ○ |
| Worked from home or stayed home from school | ○ | ○ | ○ |
| Covered your coughs and sneezes | ○ | ○ | ○ |
| Avoided going to public places, such as bars or restaurants | ○ | ○ | ○ |

*Note*. COVID-19 = coronavirus disease.

# Appendix E

Full models for logistic regressions on health-related math problem accuracy and strategy reports.

**Table E1**

*Logistic Regression Effects on Health Decision Making Accuracy for Post-test Problem 1*

| Model | Predictor | *b* (*SE*) | *Z* | *p* | *OR* [95% CI of *OR*] |
|---|---|---|---|---|---|
| Model 1 | Intercept | 0.11 (0.10) | 1.11 | .265 | 1.12 [0.92, 1.37] |
| | Condition | 1.31 (0.14) | 9.66 | <.001 | 3.71 [2.85, 4.85] |
| | Male | 0.62 (0.13) | 4.64 | <.001 | 1.85 [1.43, 2.41] |
| Model 1 versus null model, $\chi^2(2) = 117.08$, $p < .001$ | | | | | |
| Model 2 | Intercept | −0.57 (0.69) | −0.84 | .403 | 0.56 [0.15, 2.17] |
| | Condition | 1.42 (0.15) | 9.56 | <.001 | 4.14 [3.11, 5.57] |
| | Male | 0.31 (0.16) | 1.96 | .049 | 1.37 [1.00, 1.88] |
| | Baseline health decision making accuracy | 1.19 (0.16) | 7.20 | <.001 | 3.27 [2.38, 4.55] |
| | Percent absolute error | −4.40 (0.97) | −4.54 | <.001 | 0.01 [<0.01, 0.08] |
| | Math anxiety | −0.03 (0.03) | −0.87 | .384 | 0.97 [0.91, 1.04] |
| | Math attitudes | 0.05 (0.08) | 0.64 | .524 | 1.05 [0.90, 1.24] |
| | Age | <0.01 (<0.01) | 0.14 | .889 | 1.00 [0.99, 1.01] |
| | White versus nonwhite | 0.32 (0.17) | 1.94 | .052 | 1.38 [1.00, 1.91] |
| | [a]Education (bachelors) | 0.11 (0.22) | 0.49 | .623 | 1.12 [0.72, 1.74] |
| | [a]Education (graduate degree) | <0.01 (0.27) | <0.01 | .998 | 1.00 [0.59, 1.72] |
| | [a]Education (some college or associates) | −0.06 (0.17) | −0.35 | .725 | 0.94 [0.67, 1.32] |
| | Perceived COVID-19 knowledge | 0.16 (0.11) | 1.45 | .147 | 1.17 [0.94, 1.46] |
| | Health literacy | 0.01 (0.03) | 0.23 | .820 | 1.01 [0.95, 1.07] |
| | Need for cognition | 0.05 (0.10) | 0.51 | .613 | 1.05 [0.86, 1.28] |
| Model 2 versus null model, $\chi^2(14) = 260.05$, $p < .001$ | | | | | |

*Note*. COVID-19 = coronavirus disease.
[a] Each level of education is compared to participants who reported having a high school education or less. Condition was coded as 1 = intervention and 0 = control.

(*Appendices continue*)

**Table E2**
*Logistic Regression Effects on Health Decision Making Accuracy for Post-Test Problem 2*

| Model | Predictor | *b* (*SE*) | *Z* | *p* | *OR* [95% CI of *OR*] |
|---|---|---|---|---|---|
| Model 1 | Intercept | −1.53 (0.12) | −12.82 | <.001 | 0.22 [0.17, 0.27] |
| | Condition | 0.35 (0.13) | 2.69 | .007 | 1.41 [1.10, 1.82] |
| | Male | 0.64 (0.13) | 4.95 | <.001 | 1.89 [1.47, 2.43] |
| Model 1 versus null model, $\chi^2(2) = 30.49$, $p < .001$ | | | | | |
| Model 2 | Intercept | −2.20 (0.74) | −2.98 | .003 | 0.11 [0.03, 0.47] |
| | Condition | 0.40 (0.14) | 2.73 | .006 | 1.48 [1.12, 1.98] |
| | Male | 0.30 (0.16) | 1.90 | .058 | 1.35 [0.99, 1.85] |
| | Baseline health decision making | 0.94 (0.15) | 6.16 | <.001 | 2.56 [1.90, 3.46] |
| | Percent absolute error (PAE) | −7.10 (1.11) | −6.41 | <.001 | <0.01 [<0.01,0.01] |
| | Math anxiety | −0.08 (0.03) | −2.37 | .018 | 0.92 [0.86, 0.99] |
| | Math attitudes | 0.11 (0.10) | 1.17 | .242 | 1.12 [0.93, 1.36] |
| | Age | 0.01 (<0.01) | 1.24 | .215 | 1.01 [1.00, 1.02] |
| | White versus nonwhite | 0.05 (0.18) | 0.25 | .801 | 1.05 [0.74, 1.49] |
| | [a]Education (bachelors) | 0.11 (0.22) | 0.49 | .624 | 1.11 [0.73, 1.70] |
| | [a]Education (graduate degree) | −0.11 (0.26) | −0.40 | .686 | 0.90 [0.54, 1.49] |
| | [a]Education (some college or associates) | −0.11 (0.20) | −0.55 | .584 | 0.90 [0.61, 1.32] |
| | Perceived COVID-19 knowledge | 0.01 (0.12) | 0.05 | .959 | 1.01 [0.80, 1.27] |
| | Health literacy | 0.05 (0.03) | 1.55 | .122 | 1.06 [0.99, 1.13] |
| | Need for cognition | 0.12 (0.10) | 1.13 | .259 | 1.12 [0.92, 1.37] |
| Model 2 versus null model, $\chi^2(14) = 255.79$, $p < .001$ | | | | | |

*Note*. COVID-19 = coronavirus disease.
[a] Each level of education is compared to participants who reported having a high school education or less. Condition was coded as 1 = intervention and 0 = control.

**Table E3**
*Logistic Regression Effects on Whole Number Bias Errors Reported in Strategy Reports for Post-Test Problem 1*

| Model | Predictor | b (*SE*) | *Z* | *p* | *OR* [95% CI] |
|---|---|---|---|---|---|
| Model 1 | Intercept | −1.32 (0.12) | −10.63 | <.001 | 0.27 [0.21, 0.34] |
| | Condition | −0.74 (0.17) | −4.48 | <.001 | 0.48 [0.34, 0.66] |
| | Male | −0.33 (0.16) | −2.03 | .042 | 0.72 [0.52, 0.99] |
| Model 1 versus null model, $\chi^2(2) = 24.01$, $p < .001$ | | | | | |
| Model 2 | Intercept | −1.19 (0.83) | −1.44 | .151 | 0.30 [0.06, 1.53] |
| | Condition | −0.90 (0.19) | −4.78 | <.001 | 0.41 [0.28, 0.59] |
| | Male | −0.04 (0.20) | −0.18 | .859 | 0.96 [0.65, 1.43] |
| | Baseline whole number bias | 1.88 (0.19) | 9.88 | <.001 | 6.53 [4.52, 9.53] |
| | Percent absolute error (PAE) | 2.93 (1.18) | 2.47 | .013 | 18.64 [1.83, 189.91] |
| | Math anxiety | 0.10 (0.04) | 2.38 | .018 | 1.10 [1.02, 1.19] |
| | Math attitudes | −0.01 (0.10) | −0.08 | .935 | 0.99 [0.81, 1.21] |
| | Age | −0.01 (0.01) | −2.17 | .030 | 0.99 [0.98, 1.00] |
| | White versus nonwhite | −0.17 (0.21) | −0.82 | .414 | 0.84 [0.57, 1.27] |
| | [a]Education (bachelors) | −0.48 (0.29) | −1.65 | .099 | 0.62 [0.35, 1.08] |
| | [a]Education (graduate degree) | −0.76 (0.40) | −1.89 | .059 | 0.47 [0.20, 0.99] |
| | [a]Education (some college or associates) | −0.19 (0.21) | −0.89 | .372 | 0.83 [0.55, 1.25] |
| | Perceived COVID-19 knowledge | −0.16 (0.14) | −1.11 | .269 | 0.86 [0.65, 1.13] |
| | Health literacy | −0.02 (0.04) | −0.53 | .599 | 0.98 [0.91, 1.05] |
| | Need for cognition | −0.11 (0.13) | −0.84 | .402 | 0.90 [0.69, 1.16] |
| Model 2 versus null model, $\chi^2(14) = 215.44$, $p < .001$ | | | | | |

*Note*. COVID-19 = coronavirus disease.
[a] Each level of education is compared to participants who reported having a high school education or less. Condition was coded as 1 = intervention and 0 = control.

(*Appendices continue*)

**Table E4**
*Logistic Regression Effects on Whole Number Bias Errors Reported in Strategy Reports on Post-Test Problem 2*

| Model | Predictor | $b$ ($SE$) | $Z$ | $p$ | $OR$ [95% CI] |
|---|---|---|---|---|---|
| Model 1 | Intercept | −0.87 (0.11) | −7.95 | <.001 | 0.42 [0.34, 0.52] |
| | Condition | −0.29 (0.14) | −2.13 | .033 | 0.75 [0.57, 0.98] |
| | Male | −0.72 (0.14) | −5.00 | <.001 | 0.49 [0.37, 0.65] |
| Model 1 versus null model, $\chi^2(2) = 29.38$, $p < .001$ | | | | | |
| Model 2 | Intercept | −1.29 (0.73) | −1.79 | .074 | 0.27 [0.07, 1.13] |
| | Condition | −0.40 (0.16) | −2.55 | .011 | 0.67 [0.49, 0.91] |
| | Male | −0.56 (0.18) | −3.21 | .001 | 0.57 [0.40, 0.80] |
| | Baseline whole number bias | 1.51 (0.16) | 9.50 | <.001 | 4.51 [3.31, 6.17] |
| | Percent absolute error (PAE) | 3.64 (1.02) | 3.57 | <.001 | 38.20 [5.17, 284.19] |
| | Math anxiety | 0.06 (0.03) | 1.75 | .081 | 1.06 [0.99, 1.14] |
| | Math attitudes | −0.09 (0.09) | −1.06 | .289 | 0.91 [0.76, 1.08] |
| | Age | −0.01 (<.01) | −2.23 | .026 | 0.99 [0.98, 1.00] |
| | White versus nonwhite | −0.24 (0.18) | −1.33 | .182 | 0.79 [0.56, 1.12] |
| | [a]Education (bachelors) | −0.11 (0.24) | −0.47 | .642 | 0.90 [0.56, 1.42] |
| | [a]Education (graduate degree) | −0.53 (0.33) | −1.61 | .108 | 0.59 [0.30, 1.10] |
| | [a]Education (some college or associates) | −0.25 (0.18) | −1.36 | .175 | 0.78 [0.55, 1.12] |
| | Perceived COVID-19 knowledge | 0.03 (0.12) | 0.26 | .797 | 1.03 [0.81, 1.31] |
| | Health literacy | 0.01 (0.03) | 0.45 | .654 | 1.01 [0.95, 1.08] |
| | Need for cognition | −0.05 (0.11) | −0.44 | .659 | 0.95 [0.77, 1.18] |
| Model 2 versus null model, $\chi^2(14) = 220.66$, $p < .001$ | | | | | |

*Note*. COVID-19 = coronavirus disease.
[a] Each level of education is compared to participants who reported having a high school education or less. Condition was coded as 1 = intervention and 0 = control.



## WNB in Strategy Reports

For postintervention Problem 1, both when controlling for gender, $b = -0.74$, $p < .01$, $OR = 0.48$, 95% CI [0.34, 0.66], and the larger set of covariates, $b = -0.90$, $p < .01$, $OR = 0.41$, 95% CI [0.28, 0.59], those in the intervention group were less likely to report WNB errors in their strategy reports. In the simplified model, females were more likely than males to report a WNB error, $b = -0.33$, $p = .042$, $OR = 0.72$, 95% CI [0.52,0.99]. In the model including the full set of covariates, those who reported a WNB error at pretest relative to those who did not, younger participants, those with higher PAE, or those with higher math anxiety were more likely to report a whole number bias error.

For postintervention Problem 2, both when controlling for gender, $b = -0.29$, $p = .03$, $OR = 0.75$, 95% CI [0.57, 0.98], and additional covariates, $b = -0.40$, $p = .01$, $OR = 0.67$, 95% CI [0.49, 0.91], those in intervention group were less likely to report WNB errors in their strategy reports. Similar to health-related math Problem 1, the likelihood of reporting WNB errors was greater for females relative to males, in both the simplified model and the model including the full set of covariates, for those who reported a WNB error at pretest relative to those who did not, for younger participants, or those with higher PAE.

Received April 1, 2021
Revision received September 2, 2021
Accepted September 7, 2021 ■