Graphs with logarithmic axes distort lay judgments

William H. Ryan & Ellen R. K. Evers

abstract

Graphs that depict numbers of COVID-19 cases often use a linear or logarithmic scale on the y-axis. To examine the effect of scale on how the general public interprets the curves and uses that understanding to infer the urgency of the need for protective actions, we conducted a series of experiments that presented laypeople with the same data plotted on one scale or the other. We found that graphs with a logarithmic, as opposed to a linear, scale resulted in laypeople making less accurate predictions of how fast cases would increase, viewing COVID-19 as less dangerous, and expressing both less support for policy interventions and less intention to take personal actions to combat the disease. Education about the differences between linear and logarithmic graphs reduces but does not eliminate these effects. These results suggest that communications to the general public should mostly use linear graphs. When logarithmic graphs must be used, they should be presented alongside linear graphs of the same data and with guidance on how to interpret the plots.

In responding to the public’s hunger for information about the global COVID-19 epidemic, policymakers, public health experts, and journalists need to decide how to present data on the growth rate of infections in a way that lay audiences can understand easily. Even though the objective reality described by the data may be identical in different presentations, the choice of the scale used on the y-axis can greatly influence interpretations of how quickly the number of cases is changing. Being aware of the way that scale influences these interpretations is critical, because incorrect understandings can affect the public’s beliefs about the urgency of the need to take precautions to protect themselves and others.1

When presenting data on disease growth, scientists generally use either a linear or a logarithmic y-axis. A linear axis is divided into equal increments, going, say, from 100,000 cases at the first tick mark to 200,000, then 300,000, and so on. A logarithmic y-axis tracks the logarithms of values—which means that each interval on the graph corresponds to an order of magnitude increase in values rather than to a fixed increment. Thus, a jump from 100 (10^2) to 1,000 (10^3) COVID-19 cases near the bottom of an axis and the jump from 1,000 cases to 10,000 (10^4) cases a little higher up will be represented by the same vertical rise on a curve, even though the numbers of cases represented by that spatial range differs enormously. See Figure 1 for an

Figure 1. Linear & logarithmic graphs of COVID-19 cases in the United States as of May 1, 2020

Note. These graphs, from the data-visualization website https://91-divoc.com, demonstrate the difference in the appearance of curves when identical data is plotted using a linear versus a logarithmic y-axis. Titles have been added above plots and the x-axis labels have been revised for clarity. Data displays from the 91-DIVOC website have been shared virally online and used in briefings by the governors of Kentucky23 and Washington.24 The linear graph is adapted from “An Interactive Visualization of the Exponential Spread of COVID-19,” by W. Fagen-Ulmschneider, May 1, 2020 (https://91-divoc.com/pages/covid-visualization?chart=countries%show=highlight-only%data=cases&data-source=jhu&y=linear&xaxis=right#countries). CC BY 4.0. The logarithmic graph is adapted from “An Interactive Visualization of the Exponential Spread of COVID-19,” by W. Fagen-Ulmschneider, May 1, 2020 (https://91-divoc.com/pages/covid-visualization?chart=countries%show=highlight-only%data=cases&data-source=jhu&y=log&xaxis=right#countries). CC BY 4.0.
illustration of the difference between a linear and a logarithmic graph.

Logarithmic axes can be useful because they make it easier to compare exponential growth rates and to see whether case rates in different places are accelerating or decelerating similarly over time. They are also helpful when comparing countries with cases that are an order of magnitude different from one another. If one country, for example, has 10,000 cases and a second country has only 100 cases, the curve for the second country would be so tiny on a linear graph as to be unreadable. Linear graphs present case counts more directly, however, which can make them easier to grasp.

In the studies presented in this article, we asked whether the general public is more likely to misconstrue data presented on a logarithmic scale than data presented on a linear scale. We explored this possibility in part because we know of at least two cognitive processes that could contribute to such misunderstandings.

First, individuals are notoriously bad at numerical thinking in general—that is, at interpreting probabilities and other mathematical information. Numerical thinking is often measured using a test known as an objective numeracy scale, which includes math problems such as the question, “Which of the following numbers represents the biggest risk of getting a disease: 1 in 100, 1 in 1,000, or 1 in 10?” In one study, 16%–20% of participants failed to answer the simplest questions correctly. This difficulty can carry over to the interpretation of logarithmic scales, because people tend to be less familiar with logarithmic than linear scales and less able to extrapolate from the slopes they see, particularly when confronted with the exponential growth often found in data describing the numbers of people afflicted by a disease.

Second, even if individuals are able to correctly interpret logarithmic scales, they may not be motivated to do so, instead preferring to avoid numeric computations—a tendency demonstrated by an experimental tool known as the subjective numeracy scale. Studies using this instrument have found that rather than relying on active computation, people often use intuitive, so-called system 1–like, judgments, which are more likely to be biased (that is, systematically incorrect in one direction or another). Because of this intuitive processing, their judgments can be influenced by design choices. For example, identical growth data in the same type of graph can yield different inferences regarding growth rates if the aspect ratio of the graph is changed to make a curve's slope appear more or less steep. Even many experts can misunderstand graphs that use nonstandard design approaches. Such findings suggest that consumers of COVID-19-related data may tend to misjudge the severity of the pandemic when those data are presented using a logarithmic scale.

In spite of the challenges posed by logarithmic graphs, government leaders and media outlets use both types of graphs in communicating with the public about COVID-19, as seen in presentations given by the governors of Kentucky and Washington. In a further example, when we reviewed graphs in three major newspapers—the Financial Times, The New York Times, and The Wall Street Journal—we found that although linear scales were most common in articles, all three publication also used logarithmic scales. Notably, the Financial Times presented its primary COVID-19 tracker in a logarithmic format. In addition, logarithmic graphs are omnipresent in online scientific communications, which laypeople are increasingly accessing directly via social media and preprint services—such as the medical site medRxiv, which saw its number of visitors increase 100-fold between December and April.

In this article, we present four studies investigating the accuracy of predictions made by participants on the basis of each type of graph. Collectively, the studies assessed participants’ perceptions, after viewing the graphs, of the danger posed by COVID-19 and the importance of governments’ and individuals’ taking action against the spread of the disease. Overall, we found that people are reliably less accurate in their predictions of the growth rate and believe that COVID-19 is less dangerous when data are presented using a logarithmic rather than a linear scale. Additionally, this belief correlates
with less inclination to support protective behaviors recommended or mandated by governments and less inclination to adopt such behavior oneself. The Supplemental Material provides additional details on these studies’ methods and results, supplementary analyses, and information about three additional studies we conducted; also see note A.

Studies 1A & 1B: Effects on Predictions of Future Cases

Overview
In Studies 1A and 1B, we tested whether the use of logarithmic or linear scales affected the accuracy of individual’s predictions of the future rate of growth of a disease. In Study 1A, participants saw graphs displaying total COVID-19 cases in real countries and extrapolated to predict case counts in the future (see note B). Because it was possible that a country could change the criteria for counting cases during the course of a study, we also ran Study 1B: a conceptual replication in which we presented participants with hypothetical growth data for an imaginary disease and compared their predictions of future case counts with the number of cases that the equation behind the graph would predict.

Method
Participants. Participants were U.S.-based Mechanical Turk workers, who are paid to participate in research or do other online tasks. Study 1A was completed by 266 people (mean age = 38 years; 40% were female). Study 1B was completed by 403 people (mean age = 36.7 years; 37% were female).

Procedure. Both studies involved two conditions: one using a linear scale and one using a logarithmic scale. The graphs were similar to those in Figure 1. In Study 1A, each participant was shown a graph indicating case numbers up to the present for a succession of four countries (the United States first and then three others in random order). In Study 1B, participants saw a graph for one hypothetical country. In both studies, after participants viewed the graphs, they predicted the number of cases that would be seen one, three, five, and 10 days from the present for each country shown.

In Study 1B, we also measured judgments of the danger posed by COVID-19. Studies 2 and 3 also address this judgment; details of the measures used are included later in this article and in the Supplemental Material.

Results
Overall, the mean absolute error in participants’ predictions was higher in the logarithmic condition than in the linear condition in Studies 1A and 1B (see Figure 2).

The results of Studies 1A and 1B suggest that logarithmic scales tend to make laypeople’s
judgments about growth more variable than the judgments induced by linear graphs and thus less accurate. The greater variance could stem from the lower granularity in the logarithmic scale. At values above 10 cases, a given physical distance on the curve often corresponds to a greater absolute change in cases when the y-axis is logarithmic rather than linear. This feature may make it more difficult to accurately infer the number of cases higher up on the logarithmic scale.

Average variance is not the only way to measure the accuracy of participants’ interpretations. One could also ask whether presenting the data in a logarithmic format creates systematic over- or underestimation in the logarithmic group by biasing most people’s interpretations in one direction or the other. We tested this possibility but did not find a consistent pattern. We also did not find that the extent of people’s inaccuracy in predicting future case counts correlated well with their judgments of danger. It is possible that such a correlation exists but that we lacked the statistical power to detect it. Additional details on the accuracy analyses for these and the following studies can be found in the Supplemental Material.

Study 2: Influences on Beliefs & Attitudes

Overview
Of course, laypeople are not usually asked to predict the spread of a disease. What is most important for policy and for communicating with the public about COVID-19 is how the scale of linear and logarithmic graphs affects people’s perceptions of threat and need for a response. In Study 2, we looked more directly at how the choice of scale in data displays influenced participants’ judgments of not only the COVID-19 growth rate but also the threat posed by the disease and whether these judgments influenced attitudes regarding how individuals and governments should respond.

We deemed these direct measurements of views of the threat posed by COVID-19 and of the need for action to be critical because, as Study 1B suggested, it is not safe to infer such views on the basis of people’s predictions of future cases. For instance, some individuals’ judgments may be more affected by the total number of cases they foresee in the near future, whereas others may be more affected by anticipated growth rates. In addition, if some people overestimate future numbers considerably but most people underestimate the numbers slightly, the group average may suggest that people are making accurate judgments when in fact most people are underestimating the future trend and are therefore inclined to underreact to the threat of COVID-19.

Method
Participants. The study was completed by 891 U.S.-based Mechanical Turk workers (mean age = 37.9 years; 48% were female).

Procedure. The participants were presented with true disease data for the United States and then disease data for three other countries in succession in either linear or logarithmic formats and were asked questions about future case numbers in each country. Because it is possible that individuals make more accurate projections when data are embedded in a larger context, we also varied whether participants saw data of the case prevalence of only a single country or saw a country’s data on one graph that included the data for 10 additional countries.

To assess participants’ inferences about the growth rate of COVID-19 cases, we asked them to indicate how much they expected the rate to change using a scale ranging from Decrease significantly to Increase significantly. The perceived threat was measured by having participants rate COVID-19 on a scale ranging from Not at all dangerous to Extremely dangerous. To measure views of what governments should do, we had participants use a scale ranging from Disagree strongly to Agree strongly to indicate agreement with the statement that the country depicted in the graph “should ban public gatherings, close non-essential businesses, and ask all citizens to stay at home unless they are going to work or carrying out necessary errands.” We also asked participants to consider all the efforts people were taking to combat the coronavirus in each depicted country and to indicate whether
the people were exerting the right amount of effort; participants used a scale ranging from *Significantly decrease* to *Significantly increase* to indicate how they thought people should adjust their efforts. In addition to that item, we had participants indicate whether, on the basis of what they saw on the graphs, they would increase or decrease their own mask use and adherence to social distancing. (See Figure 3 for the numerical ranges of the scales.)

**Results**

Judgments of growth rate, danger, appropriate policy response, and required individual effort to combat COVID-19 were all lower when logarithmic scales were used, regardless of

Figure 3. Differences in data interpretations, attitudes, & behavioral intentions, Study 2

Note. The plots show mean responses to questions answered on the basis of viewing graphs of COVID-19 cases in multiple or single countries. Regardless of the number of countries represented on the graphs, use of a logarithmic versus a linear scale led to lower judgments of the future growth in COVID-19 cases (A), the danger posed by the disease (B), the need for policy response by governments (C), and the need for individual effort (D); it also resulted in lower intentions to use masks (E) and socially distance (F). The questions and range of replies follow. Except for B, all answers used a 7-point scale ranging from −3 (at the bottom of the axis) to 3 (at the top). A. “How do you expect the growth rate of cases, i.e. the number of new cases per day, to change in [COUNTRY]?” Range: from −3 = *Significantly fewer new cases/day* to 3 = *Significantly more new cases/day*. B. “How dangerous do you believe Coronavirus is to [COUNTRY] and its citizens?” Range: from 1 = *Not at all dangerous* to 7 = *Extremely dangerous*. C. “How much do you agree with the policy: [COUNTRY] should ban public gatherings, close non-essential businesses, and ask all citizens to stay at home unless they are going to work or carrying out necessary errands?” Range: from −3 = *Disagree strongly* to 3 = *Agree strongly*. D. “Think of all the efforts the people in [COUNTRY] may be doing to try to stop the disease, such as social distancing, wearing face masks, and avoiding non-essential travel. Based on this graph, how do you think people in [COUNTRY] should change the amount of effort they put into these actions?” Range: from −3 = *Significantly decrease effort* to 3 = *Significantly increase effort*. E and F (relating to U.S. data): “Based solely on this graph, do you see yourself wearing a mask more or less often than you do now?” (E) and “Based solely on this graph, do you see yourself being significantly more or less careful about social distancing relative to now?” (F). Range: −3 = *Significantly less often* to 3 = *Significantly more often*. 
the number of countries presented, as Figure 3 shows. Alarmingly, relative to the reports of participants in the linear condition, those who saw logarithmic graphs indicated that they would wear masks less often and be less committed to social distancing. Presenting additional countries for context reduced the perception of the danger posed to the target country in both conditions, probably because the additional countries shown all had high case counts, making the target country’s case count seem smaller in comparison.

Study 3: Education’s Ability to Reduce Prediction Errors

Some media outlets have begun educating their audience about how to correctly interpret logarithmic graphs. In Study 3, we evaluated whether this instruction brings interpretations of disease growth rates and the threat posed by COVID-19 made on the basis of logarithmic graphs in line with the interpretations made on the basis of linear graphs.

Method

Participants. This study was completed by 739 Mechanical Turk workers (mean age = 40.3 years; 50.2% were female).

Procedure. We divided the participants into two groups. One group saw linear graphs and the other group saw logarithmic graphs showing COVID-19 case data for four countries in succession. Before viewing the graphs, half of the participants in each group watched a “debiasing” video explaining how to correctly interpret linear and logarithmic graphs, and the other half of the participants—those in the control condition—watched an unrelated video of equal length about painting. The debiasing video was a 1 minute 45 second clip from the second section of a Vox Media video (available at https://youtu.be/O-3Mlj3MQ_Q?t=65.) After viewing the videos and the graphs, participants used the same scales as in Study 2 to indicate how much they expected growth rates to change and how dangerous COVID-19 seemed to them.

Results

The results are depicted in Figures 4 and 5. Informing participants about the pitfalls of logarithmic graphs reduced the difference in perceptions of growth and danger; however, those who saw logarithmic graphs still perceived lower growth and less danger than did those who saw linear graphs. That is, education helped

Figure 4. Effects of education in graph reading, Study 3

A. Intervention Effect on Growth Ratings by Graph Scale

B. Intervention Effect on Danger Ratings by Graph Scale

Note. A and B show mean responses to the same questions asked about growth and danger in Study 2. Education about how to interpret logarithmic and linear graphs reduced differences in growth and danger ratings by those who viewed logarithmic graphs but did not eliminate the gaps.
General Discussion

In the studies reported in this article, we found consistent evidence that the public formed less accurate impressions of the current COVID-19 situation when data were presented using a logarithmic scale as compared with a linear scale. Logarithmic scales also led participants to believe the pandemic was less severe and less dangerous than linear scales did. Figure 5 summarizes the growth ratings and danger perceptions reported in Studies 1B, 2, and 3. Individuals presented with logarithmic graphs were less supportive of government policies aimed at reducing the growth in cases and less likely to take individual actions such as mask wearing and social distancing.

We also explored whether a couple of interventions would minimize the misleading influences of seeing logarithmic graphs. Providing data on multiple countries did not reduce the differences between the effects of the two types of graphs, and educating participants about how to read the graphs reduced but did not eliminate the differences. We did not establish that inaccuracy in predicting future cases on the basis of seeing logarithmic plots led directly to behavior change meant to limit COVID’s spread. These two effects of reading the graphs might be independent, but researchers conducting future work should examine this issue in more detail.

We did not extensively test whether the actual growth of the outbreak could influence difference in interpretations of the two types of graphs, but we conducted one test of the possibility (as is described in Appendix Study 1 in the Supplemental Material.). To understand how the actual numbers might have an effect, consider Figure 6, which illustrates the differences between the looks of the curves depicting the first 20 days versus the first 50 days of the outbreak in the United States. Row A shows the first 20 days of the outbreak in logarithmic and linear formats, and row B adds in the next 30 days. The logarithmic and linear graphs look more similar in row A, when growth rates are still climbing, than they do in row B, when the rates have stabilized or decreased. To test whether the actual pattern of the outbreak could moderate the effects documented in this article, we ran a simple conceptual replication of the growth and danger components of Study 2, only this time we chose countries whose growth rates fit one of the two patterns in Figure 6. We
found that the logarithmic axes led to underestimation of growth and threat regardless of the slope of the actual growth.

In two of our studies, we conducted added analyses to examine whether various individual differences could influence the degree to which people struggle with logarithmic graphs. Surprisingly, we found that greater objective facility for working with numbers did not result in increased accuracy in interpreting logarithmic scales relative to linear scales. In fact, sometimes more numerate individuals fared worse than others who were less numerate. (See the Supplemental Material for a fuller discussion of the individual differences we measured.)

The skeptical reader may argue that we stacked the deck against logarithmic scales by concentrating on their influence on impressions of risk and danger rather than on their value for highlighting differences in the growth rates of cases. We argue that the choice of graph type should depend on the needs of the audience. Comparing growth rates between countries is certainly important for epidemiologists and politicians who are in a position to implement and evaluate high-level policies. For them, logarithmic graphs may be most useful. For the general public, however, graphs should be designed to enable viewers to accurately estimate risk and respond accordingly. Indeed, when we ran a test in which we expected logarithmic graphs might result in laypeople making more accurate interpretations of growth rates, we found that the claimed advantages of logarithmic graphs did not appear: for a lay audience, logarithmic graphs did not improve accuracy of comparative growth judgments. (See Appendix Study 2 in the Supplemental Material.)

**Recommendations**

Overall, our findings make it clear that officials’ and media’s decision to use either logarithmic or linear axes in graphs can influence public
response to COVID-19. When people are presented with graphs with logarithmic instead of linear axes, they make less accurate predictions of future growth; view COVID-19 as less of a threat; and, accordingly, are less supportive of governmental and individual action against COVID-19. Education can reduce these effects, but it cannot eliminate them.

Logarithmic graphs still have significant value for presenting scientific data. However, on the basis of our research, we recommend using them for the general public only when they are truly the most reasonable option. And when they are used, presenters should spend significant time explaining how to read the graphs and should supplement the logarithmic graphs with linear displays of the data.

end notes
A. Full materials, preregistrations, and data for the studies described in this article and for additional studies we conducted can be found at https://osf.io/zqu5/?view_only=3aa666d592dd2495ca508b4fa8729381a.
B. In Study 1A, actual case counts were derived from the COVID-19 data repository maintained by the Center for Systems Science and Engineering at John Hopkins University, available from https://github.com/CSSEGISandData/COVID-19.
C. It is conceivable that people who judge COVID-19 to pose a low degree of danger on the basis of seeing logarithmic graphs are more accurate in their threat assessment than are people who view the same data on linear graphs. We do not think that they are more accurate, however. When people are taught how to read logarithmic graphs, their sense of danger does not fall; rather, it rises.

supplemental material
• http://behavioralpolicy.org/publications
• Methods & Analysis

author affiliation
Ryan & Evers: University of California, Berkeley. Corresponding author’s e-mail: wryan@berkeley.edu.


