v 1.0

ACTIVITY 5

Addressing Uncertainty in Science

READING

Environmental Protection Agency researcher Megan MacDonald checks an air quality monitoring system in an industrial district in Louisville, Kentucky. Her research focuses on understanding how industrial sites contribute to air pollution emissions.

5 : ADDRESSING UNCERTAINTY IN SCIENCE

GUIDING QUESTION

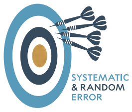
How do scientists reduce uncertainty in science?

INTRODUCTION

There can be scientific uncertainty when studying how wildfires affect health. Many studies average air quality over time (such as a year) or focus on short-term effects (such as how a week of wildfire smoke affects the number of emergency room visits). In 2024, researchers took a new approach to study the long-term effects of wildfire smoke. They looked at factors such as wildfire intensity, how often wildfires happen, and how long the smoke lasts. Rachel Morello-Frosch, a professor at the University of California at Berkeley, said, “Now that wildfires are...increasing in frequency and intensity, we can’t look at them one at a time.” The researchers found that Indigenous communities in California were exposed to 1.7 times more wildfire smoke from 2006 to 2020 than expected. This information can be used to develop new studies or to inform public policy. In this activity, you will read about one of the first large studies on the health effects of poor air quality.

The smoke from a wildfire can travel hundreds or even thousands of kilometers, depending on wind and other factors.

CONCEPTUAL TOOLS



MATERIALS LIST

FOR EACH STUDENT

- STUDENT SHEET 5.1 "Anticipation Guide: Scientific Uncertainty"
- STUDENT SHEET 5.2 "DART: Examples of Scientific Uncertainty"

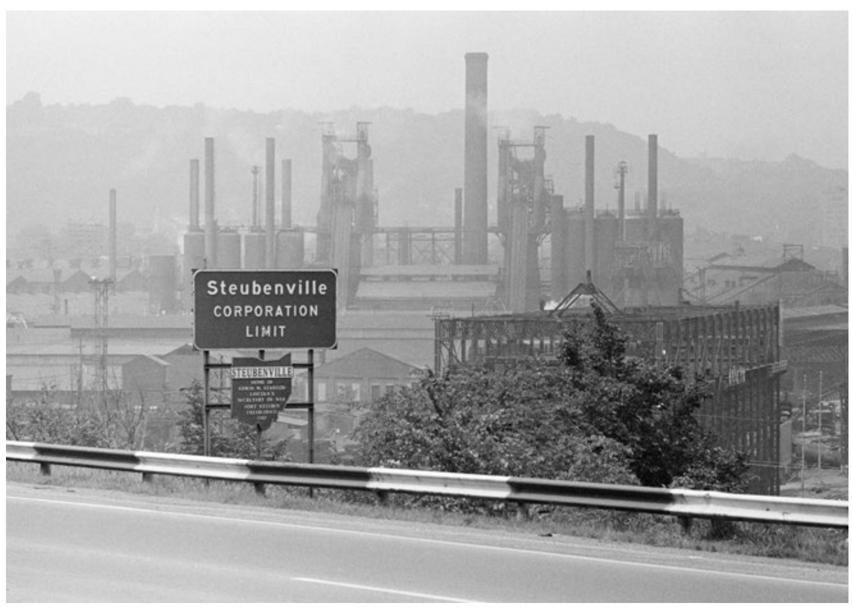
PROCEDURE

1. Complete the "Before" column on Student Sheet 5.1, "Anticipation Guide: Scientific Uncertainty," to prepare for the following reading.
2. Read the following text to learn how scientists use different methods to manage uncertainty in science.
3. As you read, fill out Student Sheet 5.2, " DART: Examples of Scientific Uncertainty." After reading, compare your ideas with another student's ideas.
4. Complete the "After" column on Student Sheet 5.1. Be sure to think about information from the reading.

READING

Air Quality and Human Health: The Harvard Six Cities Study

In the 1970s, a brown haze was a familiar sight over many cities around the world. In some places, the air quality was so poor that people had to drive with their headlights on during the daytime. Scientists began investigating the effects of air pollution on human health. They started to find links between poor air quality and increased rates of asthma, lung cancer, cardiovascular disease, and diabetes.



Steubenville, Ohio, was one of the cities studied in the Harvard Six Cities Study.

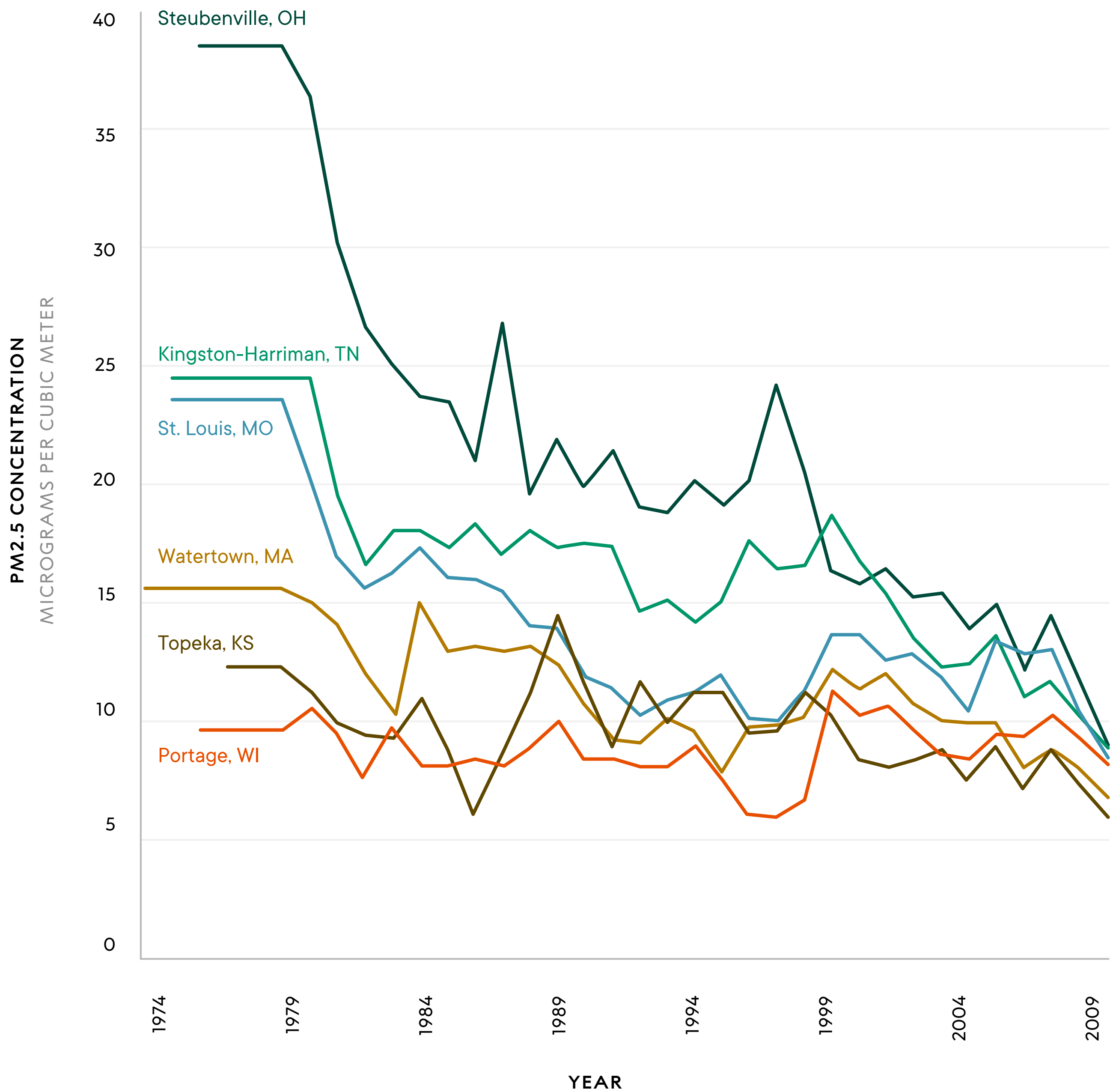
Dr. Douglas Dockery of Harvard University and his team wanted to know if poor air quality was shortening people's lives. Their study, called the Harvard Six Cities study, was published in 1993 and became one of the most important environmental health studies at the time. Between 1974 and 1989, the team collected health and lifestyle data from over 8,000 adults in six U.S. cities. They also gathered air quality data from each city. Since studying the effects of air pollution on health was challenging, the researchers needed to identify and address as many sources of scientific uncertainty as possible. This would allow them to reduce uncertainty in their conclusions once their data collection was complete.

Random Error

Most scientific research is complicated because data can have many sources of scientific uncertainty. Scientists try to identify and reduce these uncertainties, but they can never fully eliminate them. In the Six Cities study, researchers worked to minimize errors in their experimental design and measurements. For instance, air pollution levels can change from hour to hour or day to day because they are influenced by factors such as weather, wind, and industrial activity. These changes can cause random error. For example, someone using a leaf blower might carry particulate matter away from a sensor, leading to air quality readings that seem better than they really are. To reduce this kind of error, the Six Cities study averaged pollution levels over an entire year.

Scientific uncertainty from random error can also be caused by not having enough data. For instance, refer to Figure 5.1 and imagine trying to determine air quality in Steubenville, Ohio, based on just one year, such as 1996. How certain could you be that Steubenville had the highest pollution levels compared to the other cities from just one year of data? In the Six Cities study, researchers collected data from all six cities over many years. Collecting more data reduces uncertainty and makes conclusions more reliable because it increases the chance that the average of the data reflects the true value. By taking many measurements and averaging them, scientists can reduce the effects of random error. This lets them have less uncertainty in their conclusions.

FIGURE 5.1
Average Air Pollution Levels Per Year in the Six Cities , 1974–2009



Systematic Error

The Six Cities researchers also had to consider systematic error. Systematic errors can lead to inaccurate results, no matter how much data is collected or averaged. The only way to reduce a systematic error is to eliminate the source of the error. For example, if air quality sensors always measure PM2.5 levels lower than the true value, the average air pollution readings would also be too low. This would be a systematic error. Scientists regularly test their equipment and compare results from different measurement tools to reduce this type of systematic error.

Another possible source of systematic error was the placement of air quality sensors. For example, if sensors were placed in areas with cleaner air than where participants lived, the study would underestimate pollution exposure. Scientists plan for these types of errors during their study design. Sometimes they even need to redesign their experiments to avoid systematic errors.

Today, the air quality in the six cities, such as St. Louis, Missouri (pictured), has improved.

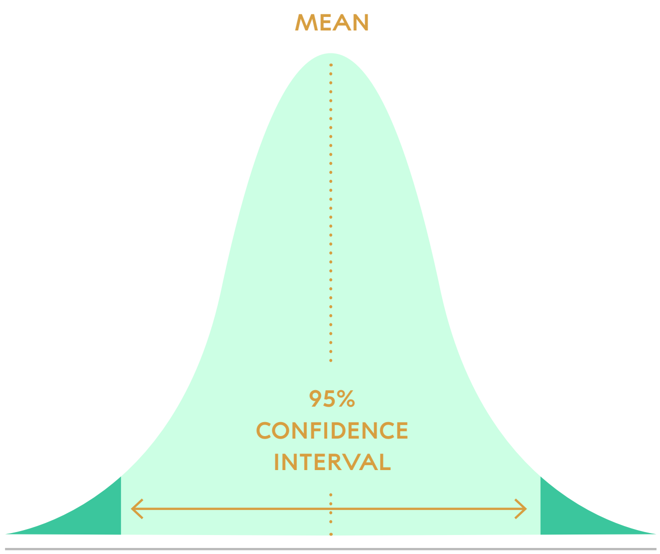
When researchers first analyzed data from the Harvard Six Cities study, they found that people in the least polluted city had a 51% higher survival rate compared to those in the most polluted city. They suspected that one of the biggest sources of systematic error came from confounds related to the health of participants. A confound is a factor that can distort or hide the relationship between two variables being investigated in a study. For example, differences in smoking habits or diet could have a stronger effect on health than air pollution. Researchers adjusted for these confounds in their data analysis. They still found a 26% difference in survival rates between people living in the least and most polluted cities. This led them to conclude that air pollution has a significant impact on life expectancy.

Communicating Scientific Uncertainty

In this unit, you have been communicating how sure you are of your data by selecting a percentage between 0 and 100. One way scientists calculate and communicate uncertainty in data is by using confidence levels and confidence intervals. In everyday life, the term *confidence* refers to feeling capable of handling a situation and achieving a goal; in science, the term *confidence* refers to a way of communicating scientific uncertainty. Scientists use a confidence level—which is a statistical measure of the probability that the true value is within a specified range. A confidence level tells you how often, in many repeated studies, the range will include the true value.

The confidence interval is calculated based on the confidence level a scientist wants to test for. A confidence interval is the range of data expected to contain the true value. Confidence intervals are probabilistic and predict how accurately your findings from a small sample apply to the larger data set. Scientists use it to estimate things such as the average height of students or the average test score in a school. Refer to Figure 5.2, which shows a confidence interval on a graph. As the sample size increases, the range of interval values narrows, meaning that you know the average with much more accuracy than you do from a smaller sample. The Six Cities study had a 95% confidence interval. That means that if the researchers were to conduct the same study 100 times, they would expect the true value to fall within this confidence interval 95 out of those 100 times.

FIGURE 5.2
A 95% Confidence Interval



Applying Science to Decision-Making

When the Six Cities study results were released, the government quickly pushed for stronger air quality standards to protect public health. However, many industries that contributed to air pollution were skeptical of the findings. They questioned whether tougher regulations were really needed. In response, many researchers reanalyzed the data and carried out additional studies to test the results. Over time, other studies confirmed the findings of the Six Cities study, including one involving over 60 million people. This additional research helped reduce scientific uncertainty in the study's conclusions.

Reducing scientific uncertainty has enabled scientists to provide clearer evidence for policymakers. The Six Cities study and the research that followed led to stricter PM2.5 air pollution laws in many countries. This helped to cut global death rates from air pollution nearly in half since 1990. Science continues to play an important role in providing data about air quality.

Face masks can be worn to reduce inhalation of particulate matter.

BUILD UNDERSTANDING

1. Researchers want to determine if reducing indoor air pollution can affect the severity of flu symptoms. The first 100 volunteers to sign up are assigned to a control group and given a device that circulates indoor air. The next 100 volunteers are assigned to an experimental group and given a device that filters indoor air pollutants before circulating the air. The researchers track the severity of flu symptoms in each group and record the number of days it takes each person to recover. At the end of the study, the researchers measure the overall health of each patient.

   Identify at least two scientific errors in the design of this study and explain how to improve the design to correct for these errors.

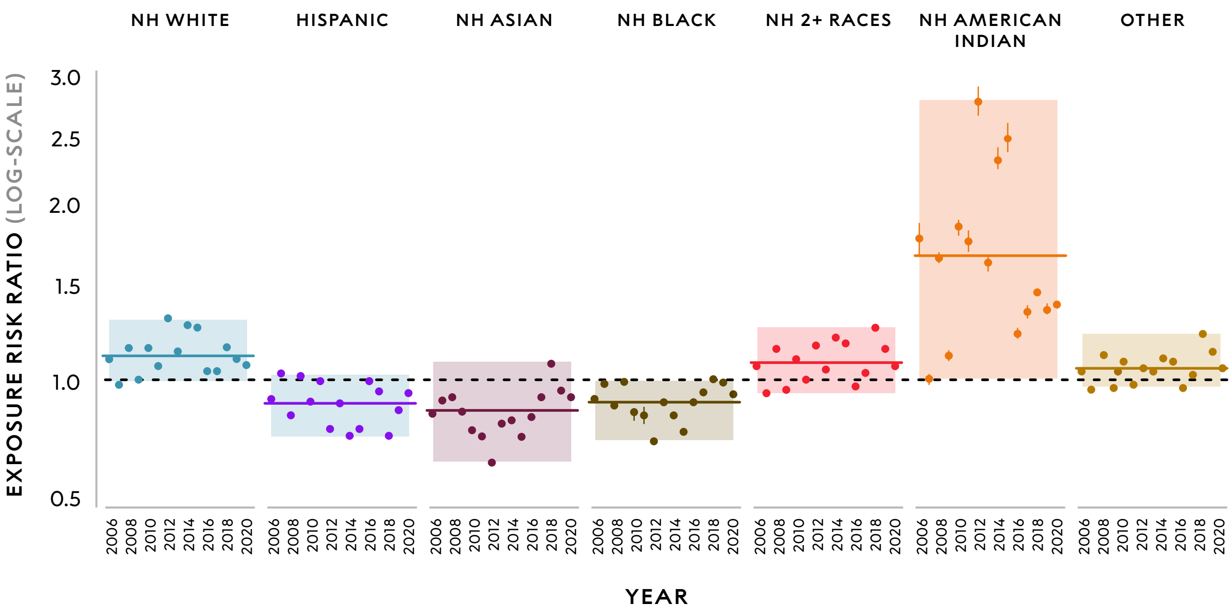
2. Figure 5.3 shows the average wildfire smoke exposure risk for different ethnic groups in California. The shaded areas of the graph represent a 95% confidence interval for each group's data set.

   a. Explain what a 95% confidence interval means in terms of this specific data.

   b. What conclusions can you make about the data?

   c. The graph for the Native American population (labeled on the graph as "Non-Hispanic American Indian") has the largest confidence interval. How does this affect the amount of uncertainty you have in the conclusions you make about the graph?

FIGURE 5.3
Wildfire Smoke Exposure, in California, 2006–2020



CONNECTIONS TO EVERYDAY LIFE

③ Johan wants to see if drinking more water will help clear up his acne. He decides to experiment by drinking 8 glasses of water every day for a month. He tracks his skin's appearance by taking pictures daily and noting the number of pimples. By the end of the month, Johan observes a lot less acne than at the beginning of the month.

The following things happened during his experiment. Explain whether each is related to random error, a systematic error due to equipment or experiment design, or a systematic error due to a confound. Explain how each might have affected his results.

- **a** Changes in weather, schedule, and diet resulted in several days when Johan drank more than or less than 8 glasses of water in a day.
- **b** Johan didn't always use the same glass when he drank the 8 glasses of water. On some days, he used a smaller glass.
- **c** After 2 weeks, Johan ran out of face soap, so he bought a different brand of face soap.

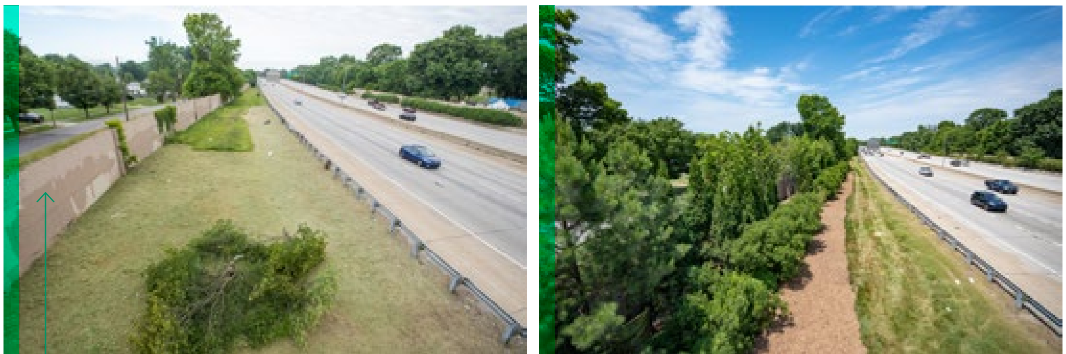
④ How would you redesign Johan's experiment from item 3 to reduce sources of scientific uncertainty in his data?

⑤ American physicist and Nobel Laureate Richard Feynman said, "When a scientist…has a hunch as to what the result is, he [or she] is uncertain. And when he [or she] is pretty…sure of what the result is going to be, he [or she] is still in some doubt. Scientific knowledge is a body of statements of varying degrees of certainty—some most unsure, some nearly sure, but none absolutely certain."

Discuss with your class how uncertainty in science can sometimes mislead people into thinking that science cannot provide trustworthy information. How would you address this issue when talking to someone who says that science does not provide true information?

EXTENSION

Are you interested in ways to improve air quality and health in your community? In 2018, a medical professor named Aruni Bhatnagar studied whether planting trees could help. He focused on neighborhoods in Louisville, Kentucky, where there were fewer trees compared to other parts of the city. These neighborhoods also had a highway running through them. Nearly 8,000 trees and shrubs were planted, and health data was collected from almost 500 residents.



Before (left) and after (right) image of an area where trees were planted along Watterson Expressway/Interstate 264 in Louisville, Kentucky, for the Green Heart Louisville Project.

In 2024, the researchers found that people living in the greener areas had 13%–20% lower levels of inflammation biomarkers, which are linked to heart disease risk. You can learn more about this project by visiting the Green Heart Louisville Project's Heal Study website. Then, brainstorm ways you could help improve and study air quality in your own community. Consider ways to manage scientific uncertainty in your design.

KEY SCIENTIFIC TERMS

confidence interval
confidence level
confound