



Lesson 12: Types of Statistical Studies

Student Outcomes

- Students distinguish between observational studies, surveys, and experiments.
- Students explain why random selection is an important consideration in observational studies and surveys and why random assignment is an important consideration in experiments.
- Students recognize when it is reasonable to generalize the results of an observational study or survey to some larger population and when it is reasonable to reach a cause-and-effect conclusion about the relationship between two variables.

Lesson Notes

A statistical study is a four-step process that begins by asking a question that can be answered with data. The next steps are to 1) collect appropriate data, 2) organize and analyze the data, and 3) arrive at a conclusion in the context of the original question. Grade 6 (Module 6, Lesson 1) introduced students to the first step in asking a statistical question while this lesson focuses on the second step on collecting data. The primary resource for the four-step process is *The GAISE Report (Guidelines in Assessment and Instruction for Statistics Education)*. You will find a free download at https://www.amstat.org/education/gaise/GAISEPreK-12_Full.pdf.

This lesson discusses the three main types of statistical studies: observational studies, surveys, and experiments. It defines each type, gives examples for each, and asks students to distinguish between them.

- An **observational study** records the values of variables for members of a sample. There are several types of observational studies. Observational studies are designed to observe subjects as they are, without any manipulation by the researcher.
- A **survey** is a type of observational study that gathers data by asking people a number of questions.
- An **experiment** assigns subjects to treatments to see what effect the treatments have on some response.

Classwork

Opening Exercise

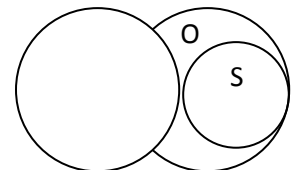
Read through the opening text as a class. Discuss each of the three types of statistical studies: observational studies, surveys, and experiments, making sure that students understand the differences between the types of studies. A visual organizer might be helpful here, especially for ELL students.

Opening Exercise

You want to know what proportion of the population likes rock music. You carefully consider three ways to conduct a study. What are the similarities and differences between the following three alternatives? Do any display clear advantages or disadvantages over the others?

Scaffolding:

A graphic organizer for statistical studies could resemble the Venn diagram below:



- a. You could pick a random sample of people and ask them the question, “Do you like rock music?” and record their answers.
- b. You could pick a random sample of people and follow them for a period of time, noting their music purchases, both in stores and online.
- c. You could pick a random sample of people, separate it into groups, and have each group listen to a different genre of music. You would collect data on the people who display an emotional response to the rock music.

A statistical study begins by asking a question that can be answered with data. The next steps are to collect appropriate data, organize and analyze it, and arrive at a conclusion in the context of the original question. This lesson focuses on the three main types of statistical studies: observational studies, surveys, and experiments. The objective of an observational study and a survey is to learn about characteristics of some population, so the data should be collected in a way that would result in a representative sample. This speaks to the importance of random selection of subjects for the study. The objective of an experiment is to answer such questions as, “What is the effect of treatments on a response variable?” Data in an experiment need to be collected in a way that does not favor one treatment over another. This demonstrates the importance of random assignment of subjects in the study to the treatments.

An observational study is one in which the values of one or more variables are observed with no attempt to affect the outcomes. One kind of observational study is a survey. A survey requires asking a group of people to respond to one or more questions. (A poll is one example of a survey.) An experiment differs from an observational study: In an experiment, subjects are assigned to treatments for the purpose of seeing what effect the treatment has on some response while an observational study makes no attempt to affect the outcomes, i.e., no treatment is given. Note that subjects could be people, animals, or any set of items that produce variability in their responses. Here is an example of an observational study: In a random sample of students, it was observed that those students who played a musical instrument had better grades than those who did not play a musical instrument. In an experiment, a group of students who do not currently play a musical instrument would be assigned at random to having to play a musical instrument or not having to play a musical instrument for a certain period of time. Then, at the end of the period of time we would compare academic performance.

Classify each of the three study methods about rock music as an observational study, a survey or an experiment.

Example 1 (7–9 minutes): Survey

Discuss each part of the example as a class.

A survey asks people to respond to one or more questions. First graders have no doubt done this in asking their classmates, “What is your favorite color?” But constructing clear questions may not be so easy. A poorly worded question may confuse the person answering it.

For example, take the question, “Do you like your school’s cafeteria food?”. How do you answer if you bring your own lunch? Or, what if you like the salad but don’t like the turkey sandwich? Assuming that the person answering eats the cafeteria food, a better survey approach would be to say: “For each of the following items offered by our cafeteria, check the appropriate box.”

Example 1: Survey

Item	I like the item.	I don’t like the item.	I have never tried the item.
Salad			
Veg pizza			
Turkey sandwich			
Raspberry tea			

- a. It is easy to determine if a study is a survey. A survey asks people to respond to questions. But surveys can be flawed in several ways. Questions may be confusing. For example, consider the following question:

What kind of computer do you own? (Circle one) Mac IBM-PC

How do you answer that question if you don't own a computer? How do you answer that question if you own a different brand? A better question would be:

Do you own a computer? (Circle one) Yes No

If you answered yes, what brand of computer is it? _____”

Now consider the question, “Do you like your school's cafeteria food?”

Re-write the question in a better form. Keep in mind that not all students may use the school's cafeteria, and even if they do, there may be some foods that they like and some that they don't like.

A survey should not be generalized beyond the set of people who responded unless participants are randomly selected from some identifiable group. How to select a random sample was covered in Grade 7, but you may need to review how to use a random-number table with your students.

- b. Something else to consider with surveys is how survey participants are chosen. If the purpose of the survey is to learn about some population, ideally participants would be randomly selected from the population of interest. If people are not randomly selected, misleading conclusions from the survey data may be drawn. There are many famous examples of this. Perhaps the most famous case was in 1936 when *The Literary Digest* magazine predicted that Alf Landon would beat incumbent President Franklin Delano Roosevelt by 370 electoral votes to 161. Roosevelt won 523 to 8.

Ten million questionnaires were sent to prospective voters (selected from the magazine's subscription list, automobile registration lists, phone lists, and club membership lists), and over two million questionnaires were returned. Surely such a large sample should represent the whole population. How could *The Literary Digest* prediction be so far off the mark?

The sample was biased toward the wealthy, and in 1936, during the heart of the Great Depression, the wealthy were not representative of the whole population of voters. Another difficulty was that not everyone who received the survey chose to return it. That resulted in what is called voluntary response bias.

Scaffolding:

Nobody appreciated the power of random sampling more than George Gallup who developed the Gallup Poll. Have your students do research on how Gallup got started in the polling business in 1936. How did he correctly predict the 1936 election while *The Literary Digest* was so far off?

- c. Write or say to your neighbor two things that are important about surveys.

Example 2 (7–9 minutes): Observational Study

Work through each part of the example as a class.

A primary difference between an observational study and an experiment is that an experiment imposes treatments on subjects to see what effect those treatments might have on their responses. Observational studies do not impose treatments on subjects. There is no attempt to influence the responses. We sometimes encounter situations in which an observational study needs to be done because an experiment cannot be performed. For example, it would be unethical (to say the least) to impose a dangerous treatment on people, such as assigning them to smoke a pack of cigarettes per day or exposing them purposely to asbestos.

A word of warning: There are many examples of famous unethical experiments that are incredibly disturbing. So, if you assign your students to find and report on an unethical experiment, you may want to provide them with a pre-screened list rather than letting them loose on the Internet.

Example 2: Observational Study

- a. An observational study records the values of variables for members of a sample but does not attempt to influence the responses. For example, researchers investigated the link between use of cell phones and brain cancer. There are two variables in this study: One is the extent of cell-phone usage, and the second is whether a person has brain cancer. Both variables were measured for a group of people. This is an observational study. There was no attempt to influence peoples' cell-phone usage to see if different levels of usage made any difference in whether or not a person developed brain cancer.

Why would studying any relationship between asbestos exposure and lung cancer be an observational study and not an experiment?

Random sampling of subjects is important in observational studies and surveys in order to eliminate bias so that sample results may be generalized to the population from which the sample was taken.

- b. In an observational study (just as in surveys), the people or objects to be observed would ideally be selected at random from the population of interest. This would eliminate bias and make it possible to generalize from a sample to a population. For example, to determine if the potato chips made in a factory contain the desired amount of salt, a sample of chips would be selected randomly so that the sample can be considered to be representative of the population of chips.

Discuss how a random sample of 100 chips might be selected from a conveyor belt of chips.

MP.3

Cause-and-effect is a very important issue, and you should allow ample time to discuss this issue and have your students critique several examples. Students might find some examples to be humorous. For example, there is a strong positive relationship between the number of television sets in a country and the life expectancies of its residents. Could one conclude that to increase the life expectancies of the residents of some country, they need only to increase the number of television sets in that country? Surely not. But why not? A "rich" country would have many television sets but would also have access to health care and fresh water which provide a better quality of life and, hence, a higher life expectancy. On the other hand, a poor country would have fewer television sets but would also have less access to health care and fresh water, leading to a lower life expectancy. The economic status of a country is what explains the relationship. Economic status is a lurking variable.

A lurking variable is one that causes two variables to have a high relationship even though there is no real direct relationship between the two variables.

Ice cream sales and number of drowning accidents are positively related. Therefore, can it be concluded that ice cream sales cause drowning accidents? The lurking variable here is time of year. More ice cream is sold in the summer than in the winter, and there is more opportunity for drowning in the summer than winter. Internet research may provide some more examples for study.

MP.2

Often people are interested in knowing what causes something to happen. The main disadvantage of an observational study is that a cause-and-effect conclusion cannot be drawn from any relationship observed in such a study.

Understanding this disadvantage is important for students to identify as they analyze statistical studies. For example, an observational study indicating that extensive use of a cell phone is indeed linked with brain cancer does not mean that extensive cell-phone usage causes brain cancer because there may be many other variables related to cell-phone usage that may also relate to brain cancer. For example, maybe people who are heavy cell phone users tend to also have very stressful jobs, and stress may be a factor contributing to poor health. An experiment must be done in order to establish causality.

- c. Suppose that an observational study establishes a link between asbestos exposure and lung cancer. Based on that finding, can we conclude that asbestos exposure causes lung cancer? Why or why not?

Answers will vary. Sample response – No, it is possible there are other variables that haven't been examined, such as whether or not the subject smokes cigarettes.

- d. Write or say to your neighbor two things that are important about observational studies.

Example 3 (5-7 minutes): Experiment

Discuss each part of the example as a class. Make sure that students understand why the study described in part (a) is an experiment and that they can identify the explanatory variable and the response variable.

The cause-and-effect difficulty discussed in observational studies can be rectified with an experiment. Being able to vary the treatment to see what the response is in each case enables a cause-and-effect conclusion to be made as long as the treatment groups are comparable. For instance, in the radish seedlings example, by having three fixed treatment levels of light/dark and having measured the lengths of germinated seedlings, we can conclude that complete darkness produces the longest lengths. (See *The GAISE Report*, pgs. 75–79.)

Example 3: Experiment

- a. An experiment imposes treatments to see the effect of the treatments on some response. Suppose that an observational study indicated that a certain type of tree did not have as much termite damage as other trees. Researchers wondered if resin from the tree was toxic to termites. They decided to do an experiment where they exposed some termites to the resin and others to plain water and recorded whether the termites survived. The explanatory variable (treatment variable) is the exposure type (resin, plain water), and the response variable is whether or not the termite survived. We know this is an experiment because the researchers imposed a treatment (exposure type) on the subjects (termites).

Is the following an observational study or an experiment? Why? If it is an experiment, identify the treatment variable and the response variable. If it is an observational study, identify the population of interest.

A study was done to answer the question, "What is the effect of different durations of light and dark on the growth of radish seedlings?" Three similar growth chambers (plastic bags) were created in which 30 seeds randomly chosen from a package were placed in each chamber. One chamber was randomly selected and placed in 24 hours of light, another for 12 hours of light and 12 hours of darkness, and a third for 24 hours of darkness. After three days, researchers measured and recorded the lengths of radish seedlings for the germinating seeds.

Recall that random sampling of subjects was done in observational studies and surveys in order to obtain a sample that is representative of a population. In experiments, subjects may not have been chosen randomly from a population but must be randomly assigned to treatments to create comparable treatment groups.

Note: There are more experimental designs than the one discussed in this lesson in which the treatment groups were independent of each other. For example, the seeds that were in the “24 hours of light” treatment were assigned to one and only one treatment. Clearly, they could not also be subjected to another treatment. But there are studies in which subjects undergo all of the treatments. The advantage of this type of design is to eliminate the effects that extraneous variables may have on the response. For example, consider a study investigating whether smelling roses improves student performance. One treatment would be to have a student wear an unscented mask and be timed on how fast he or she completes a pencil and paper maze. Another would be to time how fast a student wearing a rose-scented mask completes a pencil and paper maze. To eliminate the effect of extraneous variables (e.g., ability to do mazes, gender and age), having each student in the study do both treatments is a better design than assigning different students to the treatments, which would expose students to only one treatment each. Such a design is called a repeated measures design. Randomizing the order in which the two masks are implemented counterbalances a potential performance effect. Half of the students would do the experiment with the unscented mask first, and the other half would do the experiment with the scented mask first. The idea is that the performance effect would be about the same in each ordering and, hence, not affect the overall conclusion.

- b. In an experiment, random assignment of subjects to treatments is done to create comparable treatment groups. For example, a university biologist wants to compare the effects of two weed killers on pansies. She chooses 24 plants. If she applies weed killer *A* to the 12 healthiest plants and *B* to the remaining 12 plants, she won't know which plants died due to the type of weed killer used and which plants subjected to weed killer *B* were already on their “last legs.” Randomly selecting twelve plants to receive weed killer *A* and then assigning the rest to *B* would help ensure that the plants in each group are fairly similar.

How might the biologist go about randomly assigning 12 plants from the 24 candidates to receive weed killer *A*? Could she be sure to get exactly 12 plants assigned to weed killer *A* and 12 plants to weed killer *B* by tossing a fair coin for each plant and assigning “heads up” plants to weed killer *A* and “tails up” to weed killer *B*? If not, suggest a method that you would use.

- c. Write or say to your neighbor two things that are important about experiments.

Exercises 1–3 (20 minutes)

Have students complete the exercises with a partner or small group. Then discuss answers as a class.

Exercises 1–3

- For each of the following study descriptions, identify whether the study is a survey, observational study, or experiment, and give a reason for your answer. For observational studies, identify the population of interest. For experiments, identify the treatment and response variables.
 - A study investigated whether boys are quicker at learning video games than girls. Twenty randomly selected boys and twenty randomly selected girls played a video game that they had never played before. The time it took them to reach a certain level of expertise was recorded.

Observational study. The children were observed and no treatment was administered to them. The population of interest was all boys and girls who would play the video game that they had never played before. The study was to see who is quicker at achieving a certain level of expertise.

- b. As your statistics project, you collect data by posting five questions on poster board around your classroom and recording how your classmates respond to them.

Survey. Questions were asked of classmates. Your students may question whether this is really a sample survey since all classmates participated. They were not randomly chosen. Good point. Full marks to those who suggest that it is a census.

- c. A professional sports team traded its best player. The local television station wanted to find out what the fans thought of the trade. At the beginning of the evening news program, they asked viewers to call one number if they favored the trade and a different number if they were opposed to the trade. At the end of the news program they announced that 53.7% of callers favored the trade.

Survey. Such a sampling technique is called voluntary response. It is a very poor way of gathering data since subjects are clearly not randomly chosen.

- d. The local Department of Transportation is responsible for maintaining lane and edge lines on its paved roads. There are two new paint products on the market. Twenty comparable stretches of road are identified. Paint *A* is randomly assigned to ten of the stretches of road and paint *B* to the other ten. The department finds that paint *B* lasts longer.

Experiment. The stretches of road that are the “subjects” in this study are randomly assigned to the treatments (the two types of paint). The response variable is the longevity of the paint. Students may note that other factors could affect the experiment like weather, traffic, or other lurking variables.

- e. The National Highway Traffic Safety Administration conducts annual studies on drivers’ seatbelt use at a random selection of roadway sites in each state in the United States. To determine if seatbelt usage has increased, data are analyzed over two successive years.

Observational study. No treatment is administered. The population of interest was drivers observed at various randomly selected roadway sites. The study was to observe seatbelt usage over a two-year period. (See Making Sense of Statistical Studies, pages 60–67.)

- f. People should brush their teeth at least twice a day for at least two to three minutes with each brushing. For a statistics class project, you ask a random number of students at your school questions concerning their tooth-brushing activities.

Survey. (See Making Sense of Statistical Studies, pages 93–101.)

- g. A study determines whether taking aspirin regularly helps to prevent heart attacks. A large group of male physicians of comparable health were randomly assigned equally to taking an aspirin every second day or to taking a placebo. After several years, the proportion of the males who had suffered heart attacks in each group was compared.

Experiment. The male physicians were randomly assigned to one of two treatments. The treatment variable is aspirin or no aspirin administered. The response variable is whether or not the subject suffered a heart attack. Note that this is a well-known experiment called the Physicians’ Health Study. It has its own extensive website: <http://phs.bwh.harvard.edu/>.

2. For the following, is the stated conclusion reasonable? Why or why not?

A study found a positive relationship between the happiness of elderly people and the number of pets they have. Therefore, having more pets causes elderly people to be happier.

No. This is an observational study. No treatment was administered to the elderly. Note: You may ask your students to create the scenario as an experiment. For example, 40 elderly people were found who had the same degree of happiness on some "happiness scale." Ten of them were randomly assigned to no pets, ten to one pet, ten to two pets, and ten to three pets. Six months later, the happiness measure was taken on the people. If, for example, there was an increase in happiness as the number of pets increased, then the conclusion that having more pets causes elderly people to become happier would be reasonable.

3. A researcher wanted to find out whether higher levels of a certain drug given to experimental rats would decrease the time it took them to complete a given maze to find food.

- a. Why would the researcher have to carry out an experiment rather than an observational study?

To make a cause-and-effect conclusion requires an experiment to be done. Cause-and-effect cannot be concluded from an observational study or survey.

- b. Describe an experiment that the researcher might carry out based on 30 comparable rats and three dosage levels: 0 mg, 1 mg, and 2 mg.

The researcher should randomly assign ten rats to dosage level 0 (control), ten rats to dosage level 1 mg, and ten rats to dosage level 2 mg. The researcher should then record the time it takes each rat to complete the given maze to find food.

Closing (2 minutes)

Ask students to summarize the main ideas of the lesson in writing or with a neighbor. Use this opportunity to informally assess comprehension of the lesson. The Lesson Summary below offers some important ideas that should be included.

Lesson Summary

- There are three major types of statistical studies: observational studies, surveys, and experiments.
 - An *observational study* records the values of variables for members of a sample.
 - A *survey* is a type of observational study that gathers data by asking people a number of questions.
 - An *experiment* assigns subjects to treatments for the purpose of seeing what effect the treatments have on some response.
- To avoid bias in observational studies and surveys, it is important to select subjects randomly.
- Cause and effect conclusions cannot be made in observational studies or surveys.
- In an experiment, it is important to assign subjects to treatments randomly in order to make cause-and-effect conclusions.

Exit Ticket (5 minutes)

Exit Ticket Sample Solutions

Is the following an observational study or an experiment? Explain your answer.

Also, if it is an experiment, then identify the treatment variable and the response variable in the context of the problem. If it is an observational study, identify the population of interest.

1. A study is done to see how high soda will erupt when mint candies are dropped into two-liter bottles of soda. You want to compare using one mint candy, five mint candies, and 10 mint candies. You design a cylindrical mechanism by which the desired number of mint candies all at once. You have 15 bottles of soda to use. You randomly assign five bottles to drop one candy into, five to drop five candies into, and five to drop 10 candies into. For each bottle, you record the height of the eruption created after the candies are dropped into it.

Experiment. There are three treatments: one, five, and ten mint candies that constitute the treatment variable. The bottles were randomly assigned to the treatments. The response variable is the height the soda erupts.

2. You want to see if fifth-grade boys or fifth-grade girls are faster at solving Ken-Ken puzzles. You randomly select twenty fifth-grade boys and twenty fifth-grade girls from schools in your district. You time and record how long it takes each student to solve the same Ken-Ken puzzle correctly.

Observational study. Students are randomly selected from the population of fifth-grade boys and girls. The population of interest is the fifth-grade boys and girls in your school district. The study is to observe whether boys or girls are faster at solving Ken-Ken puzzles. No treatment was administered to the students.

Problem Set Sample Solutions

Use this space to describe any specific details about the problem set for teacher reference.

1. State if the following is an observational study, survey, or experiment, and give a reason for your answer.

Linda wanted to know if it is easier for students to memorize a list of common three-letter words (such as fly, pen, red, ...) than a list of three-letter nonsense words (such as vir, zop, twq, ...). She randomly selected 28 students from all tenth-graders in her district. She put 14 blue and 14 red chips in a jar, and without looking each student chose a chip. Those with red chips were given the list of common words; those with blue chips were given the list of nonsense words. She gave all students one minute to memorize their list. After the minute, she collected the lists and asked them to write down all the words that they could remember. She recorded the number of correct words recalled.

Experiment. The students were randomly assigned to memorize a list of common words or a list of nonsense words. There were two treatments—common words and nonsense words. The response variable is the number of correct words recalled.

2. State if the following is an observational study, survey, or experiment, and give a reason for your answer.

Ken wants to compare how many hours a week that sixth graders spend doing mathematics homework to how many hours a week that eleventh graders spend doing mathematics homework. He randomly selects ten sixth graders and ten eleventh graders and records how many hours each student spent on mathematics homework in a certain week.

Observational study. Students were randomly selected from the two populations of interest. But no treatment was administered. The population of interest is sixth graders and eleventh graders. The study was to observe the number of hours spent on mathematics homework in a certain week.

3. Suppose that in your health class you read two studies on the relationship between eating breakfast and success in school for elementary school children. Both studies concluded that eating breakfast causes elementary school children to be successful in school.

- a. Suppose that one of the studies was an observational study. Describe how you would recognize that they had conducted an observational study. Were the researchers correct in their causal conclusion?

Possible Answer: The researchers had gone to several elementary schools in a district and randomly sampled many students. They determined whether or not the students ate breakfast on a regular basis and had the district provide the students' academic performance (names of students were withheld). Their findings were not causal since the study was an observational one. The students were not randomly assigned to a treatment (eating or not eating breakfast).

- b. Suppose that one of the studies was an experiment. Describe how you would recognize that they had conducted an experiment. Were the researchers correct in their causal conclusion?

Possible Answer: The researchers had found 100 students of comparable abilities. They randomly assigned 50 of them to eat breakfast for the academic year and the other 50 to not eat breakfast. After a year, academic achievement was determined for each student. The conclusion was causal since random assignment constituted an experiment. (Note that there is an ethical issue involved in the treatment of keeping breakfast from one group of students.)

4. Data from a random sample of 50 students in a school district showed a positive relationship between reading score on a standardized reading exam and shoe size. Can it be concluded that having bigger feet causes one to have a higher reading score? Explain your answer.

No, this is an observational study. The main lurking variable is that age is the link between reading score and shoe size.

Use the following scenarios for Problems 5–7.

- A. Researchers want to determine if there is a relationship between whether or not a woman smoked during pregnancy and the birth weight of her baby. Researchers examined records for the past five years at a large hospital.
- B. A large high school wants to know the proportion of students who currently use illegal drugs. Uniformed police officers asked a random sample of 200 students about their drug use.
- C. A company develops a new dog food. The company wants to know if dogs would prefer its new food over the competition's dog food. One hundred dogs, who were food-deprived overnight, were given equal amounts of the two dog foods: the new food vs. competitor's food. The proportion of dogs preferring the new food was recorded.

5. Which scenario above describes an experiment? Explain why.

Scenario C is the experiment. The dogs were given a treatment, the choice between the new food and the competitor's food, and their preferences were recorded.

6. Which scenario describes a survey? Will the results of the survey be accurate? Why or why not?

Scenario B is the survey. No, the results would likely not be accurate because students would not answer honestly about drug usage to a uniformed officer.

7. The remaining scenario is an observational study. Is it possible to perform an experiment to determine if a relationship exists? Why or why not?

It is not possible to perform an experiment to determine if a relationship exists between mothers' smoking habits and the birth weights of babies because mothers would be randomly assigned to smoke or not smoke during pregnancy. It is unethical, and thus illegal, to tell a person that they MUST smoke.