Chapter 4: Designing Studies

1. A high-school principal discovers a web site that invites students to “Rate Your Teacher” one a 5-point scale in several categories. He thinks this might help him decide who his best teacher are, but the school’s statistics teacher tells him to ignore any information from the site. Why?

   A. The principle really should conduct an experiment to gather information on his teachers.
   AR. Incorrect. There is really no “treatment” to which he can subject his teachers (or students) that will provide him with better information about student perception of his teachers.
   *B. Only students with particularly strong opinions—and most likely, negative opinions—are likely to take the trouble to evaluate teachers on the site.
   BR. Correct. This is an example of voluntary response sample!
   C. Using a 5-point scale does not provide accurate enough information.
   CR. Incorrect. There is nothing wrong with using a 5-point scale. Think about the composition of the sample!

2. Researchers in Britain randomly divided a large number of premature babies into three groups. One received donated breast milk, one received infant formula made for premature babies, and the third received regular infant formula. Each diet was used for one month as a sole food or as a supplement to mother’s milk. Sixteen years later, the children returned and had their blood pressure measured. It was found that diastolic and systolic blood pressure both tended to be lower in the children who were fed breast milk than in the children who were fed formula. This study is an example of

   *A. an experiment.
   AR. Correct. Specific conditions (the types of diet) were imposed on the subjects (the preterm babies) and the results (blood pressure) were measured and compared.
   B. an observational study.
   BR. Incorrect. The researchers did not merely observe. They actively imposed conditions (the types of diet) on the subjects (the pre-term babies).
   C. a census.
   CR. Incorrect. The study obviously does not look at all pre-term babies.

3. A researcher in early childhood education believes that kindergarten-age children are more receptive to help from a female teacher than from a male teacher. From a list of kindergarten teachers in the state, the researcher randomly samples four classes with male teachers and four classes with female teachers. The students in the classes are interviewed, and a measure of how receptive the students in each class are to help from the teacher is determined. This study is an example of

   A. an experiment.
AR. Incorrect. For this study to be an experiment, the treatment—male or female teacher—would have to be imposed randomly on the treatment groups (the classes) after the groups were formed.
*B. an observational study.
BR. Correct. The researcher did not impose any treatment (male or female teacher) on each class. Note that confounding variables could be a problem as a result. In many schools, parents have some input in choosing their child’s teacher. When a choice of teacher is available, children might be put into a class with a teacher of the gender they would be more responsive to, which would tend to mask differences (if they existed).
C. a census.
CR. Incorrect. We are not surveying all kindergarten classes in the state. We are studying only a sample of such classes. Although this study is observational, it is not a census.

4. A study involving women aged 50 to 75 randomly assigned equal numbers of women to an exercise program (at least 45 minutes of moderate walking or riding an exercise bike five times a week) and to a stretching program (15 to 30 minutes of stretching three times a week, under the supervision of an exercise physiologist). It was found that a higher percentage of women in the exercise group reported improved sleep than did women in the stretching group. This study is an example of

*A. an experiment, but not a double-blind experiment.
AR. Correct. Treatments (the two programs) are being imposed on the subjects (the women) and their responses (quality of sleep) are being measured. The study is not double-blind because the women know which program they are following.
B. a double-blind experiment.
BR. Incorrect. The study is not double-blind because the subjects (women) know which treatment (program) they are following.
C. a matched-pairs experiment.
CR. Incorrect. A matched-pairs experiment would require that we pair subjects (women) who are similar (by age, lifestyle, and so on) and then assign one woman at random to the exercise program and the other woman to the stretching program. The study was not designed this way.

5. An agricultural scientist wants to compare the effect on yield of three different methods of growing blueberries. To control for variables such as soil condition and location, he plants 30 plots on each of six different farms. On each farm, 10 of the 30 plots are assigned to each of the three treatments (growing methods). She measures and compares the marketable yield of blueberries produced by each plot. Which of the following best describes the design of this experiment?

A. a completely randomized design with three treatments
AR. Incorrect. This is a randomized block design, because before we randomly assign the units to the treatments, we divide the units into blocks.
*B. a randomized block design with six blocks and three treatments
BR. Correct. The blocks are the farms. Within each block, we assign 10 plots to each of the three treatments (the growing methods). There are three treatments in all because there are three different distinct conditions being imposed on the units (the plots).
C. a randomized block design with three blocks and six treatments
CR. Incorrect. You have interchanged the number of blocks (the farms) and the number of treatments (the growing methods).

6. Which of the following best describes the purpose of replication in an experimental design?

A. Reducing the impact of variables other than the treatment variable.
AR. Incorrect. Random assignment to groups and controls accomplish this.
B. Repeating an experiment several times to see if results are similar.
BR. Incorrect. Scientists often replicate an experiment to confirm results, but this is not what replication means in experimental design.
*C. Using many subjects to reduce the impact of variation arising from random assignment.
CR. Correct. The more subjects you have, the less likely that, for example, all the healthy people are randomly assigned to the same group.

7. A biologist is studying the effect of different nutrients and different levels of salinity (salt) in water on the growth of a certain species of fish. Ten fish are to be assigned at random to each of 12 similar tanks in a controlled environment. The biologist wants to use combinations of 2 different nutrients and 3 different salinity levels as treatments. In this experimental design, how many factors and treatments are there?

A. Two factors and 12 treatments
AR. Incorrect. Recall that “treatment” refers to a specific combination of levels for each factor.
*B. Two factors and 6 treatments
BR. Correct. There are $(2)(3) = 6$ treatments in all, each of which will be applied to 2 tanks.
C. Six factors and 2 treatments
CR. Incorrect. You are confusing factors and treatments.

8. You want to know the opinions of American high-school teachers on the issue of establishing a national proficiency test as a prerequisite for graduation from high school. You obtain a list of all high-school teachers belonging to the National Education Association (the country’s largest teachers’ union) and mail a survey to a random sample of 2500 teachers. In all, 1347 of the teachers return the survey. Which of the following statements about this situation is true?
A. Since you took a simple random sample, it is appropriate to draw conclusions on the basis of the teachers who responded.

AR. Incorrect. Because only about 62% of the teachers responded, your results may be biased.
B. To compensate for the teachers who didn’t respond to the first survey, you can take an additional random sample of 2500 different teachers and combine the results of the two surveys.
BR. Incorrect. Increasing the sample size will not compensate for the potential bias arising from non-response. If the teachers who didn’t respond the first time had some common characteristic
that is linked to a specific attitude about national proficiency tests, the resulting bias would affect both samples.

C. To reduce bias, you should make an effort to contact and survey the 1153 teachers who did not respond to the first mailing.

*CR. Correct. A response rate of 62% suggests considerable potential for bias resulting from non-response.

9. A sociologist wants to study the attitudes of American male college students toward marriage. She gives a questionnaire to 25 of the men enrolled in Sociology 101 at her college. All 25 men complete and return the questionnaire. Which of the following is the sample in this situation?

A. American male college students.

AR. Incorrect. This is the population!

*B. The 25 men who received and returned the questionnaire.

BR. Correct. The sample is the part of the population we actually observe.

C. All male students enrolled in Sociology 101 at this college.

CR. Incorrect. This is the population from which we are drawing, not the sample.

10. A researcher is interested in the cholesterol levels of adults in the city she lives in. A cholesterol-screening program is set up in the downtown area during the lunch hour. Individuals can walk in and have their cholesterol measured for no charge. In one lunch hour, 173 people use the service, and their average cholesterol level is 217.8. The sample obtained here is an example of

A. a simple random sample, since the experimenter did not know beforehand which individuals would come to the screening.

AR. Incorrect. A simple random sample requires that all samples of a certain size drawn from the population have the same chance of being selected, clearly not the case here.

B. a stratified random sample of high- and low-cholesterol individuals.

BR. Incorrect. A stratified sample of these individuals would require a list of individuals in these two categories. Then a simple random sample would have to be taken from each list. The sample was not selected this way.

*C. a convenience sample.

CR. Correct. Individuals “volunteered” for the screening. Individuals who knew their cholesterol level was fine may not have bothered to have it checked, which would cause the average from the screening to be too high. The downtown area is not equally accessible to all adults in the population. Those who work downtown are more likely to be in the sample. Other groups may not have known about the screening or may have been unable to reach the screening center, which would lead to undercoverage.

11. You would like to compare the level of mathematical knowledge among 15-year-olds in the United States and Japan. To do this, you plan to give a mathematics achievement test to random samples of 1000 15-year-olds in each of the two countries. To ensure that the samples will include individuals from all different socioeconomic groups and educational backgrounds, you
will randomly select 200 students from low-income families, 400 students from middle-income families, and 400 students from high-income families in each country. The sampling procedure being used here is

A. simple random sampling.
AR. Incorrect. This cannot be a simple random sample because there are certain possible samples of size 1000 that have no chance whatsoever of being chosen. For example, it would be impossible to select a sample of size 1000 containing more than 200 students from low-income families, because under the stated conditions of the sampling procedure, exactly 200 such students will be in the final sample.
B. voluntary response sampling.
BR. Incorrect. You are using random selection to choose the different components of the sample. A voluntary response sampling procedure allows the members of the sample to choose themselves.
*C. stratified sampling.
CR. Correct. In this case, the strata are the different socioeconomic groups. You take random samples of the stated sizes from each of the groups and combine the samples to form the final sample of size 1000. Stratified sampling is useful when we wish to take a representative sample from a population that may contain subgroups that differ fundamentally in ways that may affect their responses. Socioeconomic status certainly plays a role in educational development.

12. A poll of American adults’ opinions about efforts to reform Social Security was conducted in 2004–2005 by the AARP, the nation’s largest organization for retired people. The poll results were criticized in some quarters because they included no respondents under the age of 30, even though voters aged 18 to 29 made up 17% of the 2004 electorate. By contrast, respondents aged 60 and above made up 34% of the sample but were only 24% of the electorate. This poll is most likely subject to which of the following types of bias?

*A. undercoverage
AR. Correct. The poll does not include the opinions of a sizable segment of the electorate. Those aged 18 to 29 did not get a chance to participate.
B. nonresponse
BR. Incorrect. No specific information was given about the response rate of the poll. However, certain groups were excluded from participation in the poll. This is not the same as nonresponse, when subjects either cannot be contacted or do not choose to participate.
C. response bias
CR. Incorrect. While it is possible that conditions under which the subjects were interviewed influenced the answers they gave, there is a much more serious problem with this sample.

13. Many people believe that taking zinc lozenges reduces the severity and duration of the common cold. Sebastien decides to conduct a study at his school to explore this claim. He sends a survey to a simple random sample of 100 students, asking them to if they took zinc lozenges or not during their last cold, and how long they experienced cold symptoms. He found that students
who took zinc had a mean cold duration that was 1.2 days lower than those who did not take zinc. Which of the following statements about his study is true?

A. Because he took a random sample, Sebastien can conclude that zinc lozenges caused the reduction in cold symptoms.
AR. Incorrect. Because Sebastien only conducted a survey, not an experiment, he cannot establish cause and effect.
B. Because he only surveyed 100 students at his school, he can only draw conclusions about those 100 students.
BR. Incorrect. Sebastien took a simple random sample, which should be representation of all the students in the school.
*C. Sebastien’s study is subject to confounding.
CR. Correct. Sebastien’s survey only shows an association between taking zinc lozenges and reduced cold duration. Perhaps students who choose to take zinc exhibit other behaviors that may reduce cold duration—such as getting enough sleep and drinking extra fluids—more than students who don’t take zinc. That is, taking zinc may be confounded with other cold-reducing behaviors.

14. Which of the following is a legitimate method for taking a simple random sample of size 50 from a population of 1000 people?

A. List the people alphabetically by last name. Starting with the first person on the list, flip a coin. If the coin comes up “heads,” that person is in the sample. Repeat this process with each person on the list until you have sample 50 people.
AR. Incorrect. For a simple random sample, every possible group of 50 people must be equally likely to be selected. This methods makes it much more likely that groups containing people whose last names are near the beginning of the alphabet are chosen.
*B. Write the names of the 1000 people on slips of paper. Put the slips of paper in a (large!) hat and draw out 50 slips.
BR. Correct. For a simple random sample, every possible group of 50 people must be equally likely to be selected. This method sill produce such a sample.
C. List the people alphabetically by last name. Choose a single random number between 1 and 50 using randInt(1,20) on a calculator. If, for example, that number is 16, choose the 50 people on the list in the following positions: 16, 36, 56, 76, . . .976, 996.
CR. Incorrect. While each individual is equally likely to be selected by this method. Each group of 50 is not. It would be impossible, for example, to have a sample consisting of the first 50 people on the list.

15. To examine how people respond emotionally to social media. A social scientist asks for volunteers at a local high school to fill out a survey. The survey asks how much time the subjects spent on Facebook, Twitter, and Tumblr in the last 48 hours, and then asks a series of questions that assess the subjects level of satisfaction with their lives at the present time. The scientist divides the subject up into “high social media use” and “low social media” use and then
compares the satisfaction ratings of the two groups. Which of the following statements is true about this study?

A. We can established whether there is a relationship between social media use and satisfaction among students at this school, but we can’t establish cause and effect.
AR. Incorrect. Because the subjects volunteered for this study, there was no random sampling. We cannot generalize our results to the entire high school.
B. We can determine whether social media influences satisfaction, but we can’t generalize beyond the subjects of this study.
BR. Incorrect. There was no random assignment to the high and low social media use groups. This we cannot establish cause and effect.
*C. We cannot determine whether social media influences satisfaction, nor can we generalize our findings to the entire school.
CR. Correct. Since there was neither random sampling nor random assignment to groups, we cannot draw any useful conclusions from this study. The social scientist needs to consult a statistician before he does any more research!