CHAPTER 3

DESIGNING STUDIES AND COLLECTING DATA

S3.1. Recent research suggests that alcoholism is partly a consequence of a genetically-based deficiency that affects the pleasure and reward areas of the brain. One recent study focused on differences in the D_2 dopamine receptor genes of alcoholics and those of nonalcoholics. There are four known alleles (genetic variants) of these receptor genes: the A_1, A_2, A_3 and A_4 alleles. The A_3 and A_4 alleles are rare, whereas the A_2 allele is found in nearly 75 percent of the general population and the A_1 allele in about 25 percent of the population.

In one study, scientists extracted DNA from the brains of deceased alcoholics and also from a group of nonalcoholics. They found that 69% of the sample of 35 alcoholics had the A_1 allele and 31% had the A_2 allele. In contrast, 20% of the 35 nonalcoholics had the A_1 allele and 80% had the A_2 allele.

a. What kind of study was this? Be as specific as possible.

ANS: *A retrospective observational study.* **(5 points)**

b. A student argues that since some alcoholics had the A_2 allele, while some nonalcoholics had the A_1 allele, these results do not demonstrate a cause-effect relationship between the presence of the A_1 allele and alcoholism. Would cause-effect be established if the study showed that 100% of the alcoholics had the A_1 allele and 100% of the nonalcoholics had the A_2 allele? If you conclude that cause-effect would not be established, what, if anything, could be established? Justify your answer.

ANS: *No, since cause-effect cannot be established by an observational study.* **(5 points)** *Only association could be established.* **(5 points)**

S3.2. Many people would choose balloon angioplasty over open-heart bypass surgery for coronary artery disease, but what is the difference in long-term results? An article in **The Boston Globe** of March 6, 1997, reported on a study in which 1,829 heart disease patients allowed themselves to be randomly assigned to receive either open-heart bypass surgery or balloon angioplasty between 1989 and 1991. Their medical conditions were then to be monitored over the course of several years, and the results compared according to several measures of cardiac health.

a. This study is a controlled experiment. Tell why.

ANS: *Treatments (balloon angioplasty and open-heart bypass surgery) were assigned by the experimenters to experimental units (patients) and responses observed (the measures of cardiac health).* **(5 points)**

b. Identify the treatments and experimental units.

ANS: *Treatments: balloon angioplasty and open-heart bypass surgery.* **(5 points)** *Experimental units: patients.* **(5 points)**

c. Suggest one way that blocking might be used in this experiment to reduce variation.

ANS: *There are a number of possibilities. Here are three: block by gender or age or specific type of heart disease.* **(5 points)**

S3.3. (25 points) In a study on the effect of caffeine intake on fine motor skills, 40 subjects were graded on their performance on a series of tests. The subjects were told to report at 8 am and were required to abstain from caffeine for 24 hours before that. The first 20 subjects to arrive were given a capsule of caffeine before performing the tests. The second 20 were given a capsule of sugar before performing the tests.

a. What kind of study was this? Be as precise as possible and justify your answer.

ANS: *A controlled experiment* (10 points)

b. There is an obvious flaw in the design of this study. Tell what it is, and how it can be corrected.

ANS: *The subjects were not randomly assigned to treatment or control. The first 20 arrivals are likely to be early risers who don't need or react to caffeine differently than do the late arrivals.* (15 points) 2

S3.4. (30 points) In a recent study, 200 runners were interviewed. It was found that those who stretched had substantially more injuries than runners who didn't.

a. Was this a controlled experiment or an observational study? Why?

ANS: *An observational study, since treatments (stretching) were not assigned to sampling units (runners).* (10 points)

b. Can we conclude from this that stretching causes injuries? Why or why not?

ANS: *No, since this is not a controlled experiment, we can only conclude there is association between stretching and injuries.* (10 points)

c. If your answer to part b is no, what alternative explanation can you provide for the observed association?

ANS: *(i) Those who stretch may do so because they have pre-existing injuries, or (ii) runners who stretch may be more competitive, resulting in more injuries, or some other explanation.* (10 points)

S3.5. Veterinary researchers want to test a promising new material for pads to cushion horse shoes. The research team has 12 horses available for testing. It is desired to evaluate the performance of the pads on the front hooves, where the greatest stress is generated. The test apparatus consists of a short track with a pressure plate mounted flush to the ground. The pressure plate is large enough for the horse to land, and have the pressure recorded, for both front hooves on one run. The horses will be run at a near-constant speed on each run. The head researcher suggests that six of the horses be chosen randomly to test the new pad on both hooves, and that the remaining horses be given the old pad on both hooves.

a. Is this a controlled experiment? Justify your answer.

ANS: *Yes, it is a controlled experiment, since treatments (pad types) are imposed on experimental units (horse hooves) in order to observe a response (pressure).* **10 points)**

b. Suggest a way of conducting the study that is superior to that suggested by the head researcher. Tell why your design is superior.

ANS: *Put one pad of each type on each horse.* **(5 points)** *This kind of blocking is superior since it removes horse-to-horse variation.***(5 points)**

S3.6. The article "Study Says Rigor of High-School Course Work Is the Best Predictor of College Graduation," by Ben Gose, Chronicle of Higher Education, XLV, 39, June 4, 1999, p. A46, describes a study conducted by the U. S. Department of Education purporting to show that the rigor of a student's high school curriculum is a better predictor of whether the student will graduate from college than either test scores or high school grades. The following is a quote from the article:

Clifford Adelman, the senior research analyst who led the study, spent two years examining the transcripts and test scores of students who graduated from high school in 1982 to determine what contributes most to the completion of a bachelor's degree.

Mr. Adelman established college graduation by the age of 30 as a benchmark. He then worked backward to see which measure of preparedness for college–the student's test scores, class rank, or course work–had the strongest correlation with the earning of a bachelor's degree.

The study was reviewed by Alberto F. Cabrera, an associate professor and research associate at the Center for the Study of Higher Education at Pennsylvania State University. Mr. Cabrera says the study should pursuade high-school guidance counselors to encourage as many students as possible to take demanding courses. "The findings make a clear point that the type of courses a student takes has a long-term effect on the student's success," Mr. Cabrera says.

a. Tell as precisely as possible what type of study is described here. Justify your answer.

ANS: *This is an observational study, since it is not a controlled experiment (No treatments are imposed on experimental units; rather, attributes of the sampling units (the students) are observed from records.)* **(10 points)** *Further, it is a retrospective study, since the students are classified after the fact into graduates and non-graduates, and then factors that differ substantially between these two groups are sought.* **(10 points)**

b. Explain why the conclusion stated in Mr. Cabrera's quote is not justified.

ANS: *Cause-effect cannot be inferred from an observational study alone. It is very possible (in fact, I think more likely) that what we are observing is that good students take tough courses and good students graduate from college at higher rates. Put another way, it is problematic whether putting a poor student in honors and AP courses in high school will cause her to perform better in college.* **(10 points)**

S3.7. **(20 points)** A study directed by Dr. Michael Gaziano, a heart specialist at the Harvard Medical School-affiliated Brigham and Women's Hospital in Boston, suggests that tea may reduce the chance of a heart attack by 44 percent. Gaziano's study examined 340 men and women who had suffered heart attacks and matched them by age, sex and neighborhood with people who had never had a heart attack. It then investigated their coffee- and tea-drinking habits over the course of a year. Gaziano thinks that the reason for the reduced incidence of heart attacks among tea drinkers is that tea contains powerful amounts of flavonoids, natural substances that make blood cells less prone to clots, which can cause heart attacks. The study was adjusted for factors that could have skewed the results, such as smoking, exercise, alcohol intake and family history of heart trouble. Total calories consumed, intake of fatty foods and body mass index which compares the girth of people of different heights to determine obesity was about the same across the board.

a. What kind of a study was this? Be as precise as you can, and justify your answer.

ANS: *A retrospective observational study, since subjects are classified into groups after the fact, and causal factors sought.* **(10 points)**

b. Based on this study, is it valid to conclude that drinking tea causes a reduction in the chance of a heart attack? Why or why not?

ANS: *No, because, even though the researchers tried to adjust for extraneous factors, this is not a controlled experiment.* **(10 points)**

S3.8. A researcher conducted a study of the effectiveness of an SAT preparation course. To do so, she randomly selected a group of 100 high school juniors from a large high school. She randomly divided them into two groups of 50 each. Both groups took the SAT on the same date in the fall. Then, one group was given the SAT preparation course and the other was not. Finally, both groups took the SAT on the same date in the spring. To evaluate the results, the differences in the two SAT scores for each student were computed, and these differences were compared for differences between the two groups.

a. (10 points) Explain why this a controlled experiment.

ANS: *It is a controlled experiment since treatments (prep course/no prep course) were imposed on experimental units (students) to observe a response (difference in SAT scores).*

b. (10 points) How is blocking used in this experiment? What do you think its effect will be?

ANS: *Students are blocks in which both the before and after SAT score are observed. It should eliminate between-student variation, allowing the researcher to focus on the effect of the prep course.*

c. (10 points) Randomization is used in two ways in this experiment. Explain how it is used and why.

ANS: *Randomization is used to select the sampling units from the larger population. This will enable the results to be extrapolated to that larger population. It is also used to assign treatments to experimental units. This helps guard against unsuspected sources of bias in the two groups.*

d. (10 points) Describe how replication is used in this experiment.

ANS: *There are 100 experimental units: 50 in each group. These are replicates.*

S3.9. A study was conducted to compare the abilities of two commercial sunscreens to resist washing off in pool water. Twelve volunteers were randomly divided into two groups of 6. A prescribed amount of sunscreen 1 was applied to both arms of each subject in group 1, and the same prescribed amount of sunscreen 2 was applied to both arms of each subject in group 2. The subjects were assigned at random to the 12 pool lanes, where they simultaneously swam for 30 minutes. At the end of the time, the amount of sunscreen remaining on the arms of each subject was measured.

a. **(15 points)** Tell why this study is a controlled experiment. Justify your answer with specifics.

ANS: *It is a controlled experiment because treatments (sunscreens) were applied to experimental units (arms) to observe a response (remaining sunscreen).*

b. **(10 points)** How could blocking be used to improve the experiment? Explain why you think it might result in an improvement.

ANS: *By applying one sunscreen to one arm, and the second to the other arm, each subject would become a block. Such blocking would reduce subject-to-subject variability in terms of the response.*

S3.10. Three studies are described below. Tell as completely as possible what type of study each is. Justify your answers.

a. **(10 points)** Scientists randomly divide volunteers into two groups. In one group, each individual is given a geometric puzzle to solve, while each individual in the other group is given no stimulation.

After five minutes, magnetic resonance imaging (MRI) measurements of the individual's brain activity are obtained.

ANS: *This is a* **controlled experiment** *since treatments (puzzle or no activity) are assigned to experimental units (subjects) in order to observe a response (MRI measurements).*

b. **(10 points)** Crop yields are estimated by satellite. First the US is divided into sectors that can be viewed by the satellite in a single image. Images are obtained of a set of sectors selected by a probability sampling scheme. These images are analyzed to estimate the total US crop yield.

ANS: *This is a* **sample survey** *since a set of sampling units (sectors) is selected from the population of sectors to obtain information about that population.*

c. **(10 points)** In order to identify possible causes of tread separation on a particular model of SUV tire, quality engineers compare a sample of failed tires with a sample of unfailed tires with respect to a number of variables.

ANS: *This is a retrospective observational study, since tires are grouped by end result (failure or not) and differences in hypothesized causes sought.*

S3.11. **(10 points)** Describe how the first study in exercise S3.11 could be improved by the use of blocking. Be sure to tell how the blocking would help.

ANS: *Use each subject as a block and measure each under treatment (puzzle) and control (no stimulation). This would help reduce subject to subject variation.*

S3.12. **(10 points)** Are the results of the last study in exercise S3.12 necessarily applicable to the population of all tires of the model studied? If you conclude 'yes,' tell why. If you conclude 'no,' tell why not and indicate what must be done to make the results applicable to that population.

ANS: *They are not necessarily applicable, since the samples may not have been chosen in a way that makes it reasonable to assume they are representative of the population. The tires should be chosen according to an appropriate probability sampling scheme in order to justify extending the results to the population.*

S3.13. It is well known that anti-perspirants applied to the skin can cause skin irritation in some people. The manufacturer of a new anti-perspirant wants to test the degree of skin irritation induced by the active ingredient in its product. To do so, it recruits 20 volunteers and randomly selects 10 to test its anti-perspirant. The other 10 are assigned a formulation identical to the product, but without the active ingredient. All subjects are instructed to apply the product once per day to the armpits, and are assessed at regular intervals over a four week period for skin condition in the areas of application.

a. **(10 points)** Is this a controlled experiment? Why or why not?

ANS: *It is a controlled experiment because treatments (anti-perspirant or placebo) were applied to experimental units (subjects) to observe a response (skin condition).*

b. **(10 points)** If it is a controlled experiment, identify the treatments, experimental units, and response. If it is not a controlled experiment, identify the kind of study it is.

ANS: *(See above)*

c. **(10 points)** Assuming the budget only allows 20 subjects for this study, tell how the study can be improved by blocking.

ANS: *By having the subject apply the product to one armpit, and the placebo to the other armpit, each subject would become a block. Such blocking would reduce subject-to-subject variability and allow a more direct comparison of the product to the control.*

S3.14. In 1976 the Viking spacecraft ran a remote experiment to test for life on Mars. The experiment involved feeding nutrients to a sample of Martian soil and measuring the amount of carbon dioxide produced. The theory was that if there were microbes in the Martian soil similar to those found on earth, the nutrients would be absorbed by the microbes and then given off as carbon dioxide. The experiment yielded the kind of positive results expected if there were microbes present in the soil, but, because other tests, including one that found no trace of organic chemicals in the soil, led most scientists to conclude that the nutrient experiment had been fooled by some kind of exotic and chemically active substance in the soil.

Environmental engineer Gilbert Levin, designer of the 1976 experiment, proposed a similar but improved experiment for the Mars Surveyor '98 spacecraft.

The new experiment was based on a peculiar, unexplained feature of all known life: It likes only left-handed molecules. Most of the organic molecules on Earth are asymmetrical, and can exist in two mirror-image forms. But living organisms use only one form, called left-handed. In contrast, all known non-organic chemical reactions work equally well on either left- or right-handed molecules.

Levin's proposed experiment was to obtain two samples of Martial soil, feed one sample nutrients made up of only left-handed molecules, and the other sample the same amount of nutrients made up of only right-handed molecules and measure the carbon dioxide produced in each sample.

a. **(10 points)** Explain why each of the described experiments is a controlled experiment.
   ANS:*For both experiments treatments (nutrients) were imposed on experimental units (soil samples) to observe a response (CO_2 produced).*

b. **(10 points)** What are the experimental units, treatment(s) and response for each?
   ANS: *(See Above).*

c. **(10 points)** What possible observed outcomes do you see from the second experiment? For each possible outcome, tell what scientists can conclude.
   ANS: *Outcome 1: no difference in CO_2 produced from the two samples. Conclusion: no microbes present. Outcome 2: more CO_2 produced from the sample given the left-handed molecules. Conclusion: microbes present.*

d. **(10 points)** Explain why the proposed experiment is an improvement over the 1976 experiment.

ANS: *It uses a control and treatment, i.e., uses the method of comparison.*

S3.15. **(10 points)** A shoe manufacturer wants to compare two materials for use on the soles of athletic shoes. To do so, 100 volunteer runners are randomly assigned to two groups. Each volunteer in the first group is given a pair of athletic shoes with material 1, while each volunteer in the second group is given a pair of athletic shoes with material 2. They are asked to return the shoes after they have run 500 miles in them, so that wear of the two materials can be compared. Show how blocking can be used to give a better design for this controlled experiment, and explain why your design is better.

ANS: *Assign to each runner one shoe with material 1 and one with material 2. This will eliminate runner to runner variation.*

S3.16. **(10 points)** If a data distribution is symmetric, you know that the mean equals the median. Is the converse true: that is, if the mean of a data distribution equals its median, must the distribution be symmetric? Prove it's true or give an example to show it isn't.

ANS: *Not true. E.g.: 6 data values, -4, -2, -1, 1, 3, 3. Median=mean=0.*