Section 4.1 Experimental and Observational Studies

4.1 Cell phones

a) The response variable is a specific type of brain activity. The explanatory variable is whether or not the automated call was placed to the phone on the right ear.

b) This was an experiment because researchers controlled via randomization whether the call to a given participant would be received during the first PET scan or the second and then measured the participant's brain activity under both treatments.

4.2 High blood pressure and binge drinking

a) This was an observational study because the experimenter did not assign subjects to treatments.

b) The response variable is whether or not a subject dies from stroke or heart attack. The explanatory variable is whether the subject has high blood pressure and binge drinks even occasionally or was a teetotaler with normal blood pressure.

c) This does not prove that a combination of high blood pressure and binge drinking causes an increased risk of death by heart attack or stroke. There could be a third variable associated with both high blood pressure/binge drinking and death by heart attack or stroke.

4.3 Low-fat versus low-carb diet?

a) The response variable is weight loss (weight after 1 year minus weight before experiment). The explanatory variable is whether one is on a low-fat or low-carbohydrate diet.

b) This was an experimental study since the subjects were randomized by the researchers into one of the two groups.

c) No. This is an issue of generalizability. Because patients with heart disease or diabetes were excluded from the study, it is inappropriate to say that this applied to *everyone*.

4.4 Experiments versus observational studies

An experiment is preferred over an observational study when either is feasible because it tells us something about cause and effect. In an experiment, the researcher has more control over the levels of the explanatory variable, and is thus able to reduce the possibility of lurking variables. When subjects are randomly assigned to treatments, the groups tend to be balanced in terms of possible lurking variables. For example, if we were interested in whether the use of an online study tool increased statistics grades, we could conduct either an observational study or an experiment. In an observational study, we would ask students whether they used an available online tool (the explanatory variable) and would assess their grades (the response variable). In an experiment, we would randomly assign students either to use the online tool or not to use the online tool. If our studies supported our hypothesis that the online tool increased student grades, the experiment would be a stronger finding because random assignment likely made the groups balanced in terms of other characteristics, such as hours of studying. With the observational study, it is possible that the students who chose to use the available extra tool were the students who tended to study more to begin with, and thus, received higher grades because of their higher overall levels of studying. However, it is not always possible for researchers to carry out a study in an experimental framework. Many factors can make experimentation impossible or next to impossible. One such factor is ethical concerns. For example, one could not study the effect of smoking during pregnancy by assigning one group of pregnant women to smoke during pregnancy and the other to not smoke.

4.5 Tobacco prevention campaigns

a) According to the study, an increase in y will surprisingly increase x. Therefore, a positive association might exist between the two variables.

b) Although tobacco companies spend huge sums of money on advertising and promotional expenditures for cigarettes and smokeless tobacco, they know in advance that such campaigns have no impact on the initiation or prevalence of smoking among young people. Tobacco companies launch youth anti-smoking campaigns as public relations efforts to discourage regulation and public action to reduce smoking, such as litigation or legislation, with the goal to improve their public image and reduce opportunities for opponents to impose restraints on industry practices. Their targets, then, are not youth or parents of youth, but policymakers and the general public.

4.6 Hormone therapy and heart disease

a) It is possible that a lurking variable such as health-consciousness caused the observed difference. Perhaps women who were higher in health-consciousness had better access to the hormone-replacement drug than did those lower in health-consciousness. In this case, it is possible that health-consciousness and the better nutrition, better overall healthcare, and other health benefits that accompany it are the cause of the decreased risk of heart disease. (Other possible lurking variables are genetics and wealth.)

b) Different types of studies can lead to different findings. An observational study is more susceptible to effects of lurking variables, such as health-consciousness, so its results could easily differ from those of a controlled experiment. An experiment would randomly assign subjects to treatments – hormone-replacement drugs or not – and thus would have more balanced groups. Levels of health-consciousness and its associated benefits would be more evenly distributed between groups, and differences in the response variable would be more likely due to the explanatory variable.

c) Randomized experiments, when feasible, are preferable to observational studies because they reduce the effects of lurking variables. Randomization into the treatment groups will in theory evenly spread out the effects of the lurking variables within all the treatment groups.

4.7 Children of mothers with remitted depression

a) The explanatory variable is the health history of a mother (with/without remitted depression). The response variable is the exposure of children to sadness.

b) This study is observational, no assignment of treatments was made by the researchers; the researchers simply observed the health history of the subjects and the exposure to sadness of their children.

c) No, possible reasons some mothers had depression history such as poverty or family instability might be the true causes of their children's exposure to sadness.

4.8 Breast-cancer screening

a) This is an observational study since the women were not assigned to treatment groups.

b) The response variable is whether or not the woman died from breast cancer during the time period of study. The explanatory variables are time period during which the woman was observed (1996 to 2005 or 1986 to 1995) and whether or not she was living in a country with mammography screening.

c) The study does not prove that being offered mammography screening causes a reduction in death rates associated with breast cancer because observational studies cannot demonstrate causation. There could be other variables associated with mammography screening and breast cancer survival rates.

4.9 Experiment or observe?

a) Observational study (unethical to assign people to caffeine consumption)

b) Experiment (can assign some classes to multitasking in-class activities)

c) Observational study (can't assign students to study abroad)

4.10 Baseball under a full moon

a) The comment is based on observational data.

b) No, the Boston Brouhahas should not be concerned. It is more likely that it is mere coincidence or that there is a lurking variable affecting his observed finding.

4.11 Seat belt anecdote

Anecdotal evidence cannot be expected to be representative of the whole population. The seat belt incident might be the exception, rather than what is typical. Death rates are in fact higher for those who do not wear seat belts.

4.12 Job opportunity in New York City

No, Tina's mother findings about homicides rate in New York City should not drive her daughter to reject the offer. Data published online by the New York Police Department indicate that crime rates are showing a significant dropping trend. Moreover, some useful tips could be recommended to avoid running into trouble in New York City.

4.13 What's more to blame for obesity?

a) This is an observational study since no assignment of treatments was made by the researchers; the researchers simply observed current habits of the subjects.
b) The response variable is weight gain. The explanatory variables are exercise habits and caloric intake.
c) No, it demonstrates a correlation, but observational studies cannot demonstrate causation. An experiment would be needed to show causation.
d) Motherhood leads to less exercise, eating more, and a more sedentary lifestyle.

💻4.14 Census every 10 years?

a) Censuses are extremely costly and time consuming.
b) The census gives the government a count of the population, enabling it to make better plans for health, employment, education, etc. in the future. It also allows the government to re-apportion congressional seats.
c) Answers will vary. Examination of the form will show that of the 10 questions on the form, Question 6 pertained to gender and Question 8 pertained to race.

Section 4.2 Good and Poor Ways to Sample

4.15 Choosing officers

a) PV PS PT PA VS VT VA ST SA TA
b) Given that there are ten combinations, there is a one in ten, or ten percent chance that a particular sample of size two will be drawn.
c) The Activity Coordinator is in four of the ten combinations; thus, there is a four in ten, or forty percent chance that she/he will be chosen.

💻4.16 Simple random sample of students

Answers will vary. To use a random number generator, select the minimum value as 1, the maximum value as 60, and the number to select to two. Using a random number table, select two-digit numbers until you have obtained two distinct values between 01 and 60.

💻4.17 Auditing accounts–app

Answers will vary each time this is run. The accounts should be labeled from 01 to 60, then random two-digit numbers should be generated. Using a random number table, select the first ten two-digit numbers that fall between 01 and 60. Ignore duplicates.

💻4.18 Sampling from a directory

Using a random number table, you would select five-digit random numbers, ignoring 00000, and numbers above 50,000, as well as duplicates until 10 numbers were found. Using a random number generator, you would select the minimum value as 1, the maximum value as 50,000, and the number to select to ten. Then you would find the names associated with those ten numbers. If, for example, you selected the number 13,050, you would turn to the 131st page (which would include 13,001 to 13,100), and then select the fiftieth name. You would continue until you had ten names.

4.19 Bias due to interviewer gender

This example illustrates response bias because at least some subjects are not giving their true opinions.

4.20 Charity walk

a) This is a leading question because respondents' answers will be influenced by the claimed consequence of saving hundreds of teenagers from suicide. The organization tries to twist the research to support a specific choice.
b) This is a better way to ask the question because there is no preference for one location on the other within the question.

4.21 Instructor ratings

Comment **a.** is valid. The ratings are a volunteer sample and likely not representative of the population. Comment **b.** is invalid. Taking a simple random sample of a biased sample will not provide you with an unbiased sample.

4.22 Job trends

a) The population for this survey was employers. Theoretically, it would be all employers in the United States.
b) The number of employers that were surveyed (the sampling frame) is required to calculate the nonresponse rate.
c) Two potential sources of bias are nonresponse bias and voluntary response bias.

4.23 Gun control

a) Someone who opposes gun control would prefer to quote the statistic that says that more than 75% of Americans say no when asked "Would you favor a law giving police the power to decide who may own a firearm?".
b) Both statements would be considered leading, and could increase response bias. The first mentions "illegal gun sales" and the second states "giving police the power to decide" – both statements might sway people's responses, albeit in different directions.

4.24 Physical fitness and academic performance

a) The population consists of all the fourth and fifth grade students from a rural northeast Georgia elementary school.
b) This is an observational study as no assignment of treatments was made by researchers; the children were simply observed by the researchers.
c) Possible lurking variables could be the self-esteem level and the level of anxiety. Physical activity is linked to higher levels of self-esteem and lower levels of anxiety, which are associated with higher academic performance in the classroom.

4.25 Fracking

a) The population is all adults in the United States.
b) It is almost impossible to ask all adults in the United States their opinions. A random sample gives a reliable estimate of this value with a much smaller set of people.
c) Those who chose to respond might feel more strongly than those who did not respond.

4.26 Sexual harassment on the Internet

a) This study likely has sampling bias since the sampling method was not random. Not all Internet users are equally likely to respond to an Internet survey administrated through social media networks.
b) It is possible that few Internet users who have been engaged in Internet sexual harassment activities are unlikely to respond to the survey. This would introduce nonresponse bias into the study.
c) It is also possible that not all Internet users answer the survey question truthfully, particularly if they are fearful of getting in trouble for answering in the affirmative.

4.27 Cheating spouses and bias

a) Those who don't admit to being in an extramarital affair might spend more money to hide the fact. Therefore this estimate would be too low.
b) If people lied about how much money they spent on an affair, it seems likely that they would underestimate the amount spent, to assuage their guilt.

4.28 Drug use by athletes

a) It is likely that the reporter's study contains sampling bias due to undercoverage because it is unlikely that all athletics fans are represented by the study.
b) Since the sampling design did not consist of a random sample (participants of the study were self-selected), it is very likely that there is sampling bias due to the sampling design.
c) Response bias is probable in this study, surveyed athletics fans are likely to give subjective answers based on personal opinions rather than trustfworthy scientific evidences.

4.29 Identify the bias

a) Undercoverage occurs because not all parts of the population have representation, only those who have subscribed to this newspaper the longest.
b) One problem with sampling design is that the newspaper does not even use random sampling among its subscribers; but instead the 1000 people who have subscribed the longest.

4.29 (continued)

c) Among those who are sent a questionnaire, not all will respond. It is possible that those who respond feel more strongly about the proposal. There was a high percentage of people who did not respond.

d) The question is framed in a negative way, perhaps skewing people's responses in a negative direction – a kind of response bias.

4.30 Types of bias

a) A convenience sample, rather than a random sample, might lead to sampling bias. For example, if a researcher collected survey data on exercise at an expensive local gym, the sample would not likely reflect the general population of people who exercise.

b) Undercoverage occurs when the sampling frame (the list of subjects from which the sample is taken) does not include some part of the population. For example, a survey conducted by emailing people would omit from the sampling frame all people without email addresses. These people likely differ in important ways from those who complete the survey.

c) Response bias might occur in a survey on drug use in a local high school because students might lie to give what they think the researchers view as an acceptable response.

d) A survey about voting intentions in an upcoming election might suffer from nonresponse bias in that those who feel most strongly about the election might be more likely to respond, thus skewing the results.

Section 4.3 Good and Poor Ways to Experiment

4.31 Smoking affects lung cancer?

a) This is an experiment because subjects (students in your class) are randomly assigned to treatments (smoking a pack a day, or not ever smoking). Subjects do not choose whether or how much they smoke.

b) One practical difficulty is that it is unethical to assign students to smoke, given that it could cause lung cancer. A second is that we would have no way to ensure that our subjects do as assigned and smoke or not smoke according to the assignment. A third is that we would not have results for fifty years – too long to wait for an answer.

4.32 Never leave home without duct tape

a) The response variable was whether the wart was successfully removed, and the explanatory variable was type of treatment for removing warts. The experimental units were the 51 subjects between the ages of 3 and 22. The treatments were duct tape therapy and cryotherapy.

b) You could label the 51 patients from 01 to 51. You could then pick two-digit random numbers until you had chosen 25 numbers between 01 and 51 and thus 25 patients to assign to one treatment. The rest would be assigned to the second treatment. You would disregard 00 and any number over 52, as well as duplicates.

4.33 More duct tape

a) The response variable was whether the wart was successfully removed and the explanatory variable was the type of treatment for removing the wart (duct tape or the placebo). The experimental units were the 103 patients in the Netherlands. The treatments were duct tape therapy and the placebo.

b) The difference between the number of patients whose warts were successfully removed using the duct tape method and those using the placebo was not large enough to attribute to the treatment type. In other words, the difference in the success rates could be attributed to random variation.

4.34 Fertilizers

a) The response variable is whether or not the soil gets more fertile during the study period.

b) The explanatory variable is the treatment type: eight different organic and inorganic fertilizers.

c) The experimental units were subject subplots chosen with identical specifications and observed during the study period.

d) The treatments were the different types of organic and inorganic fertilizers.

e) The differences were large enough to support that the observed effect of organic fertilizers was due to something other than ordinary random variation.

4.35 Facebook study

a) This is an experiment because there was a control group and an experimental group and the researchers assigned the Facebook users to each group.
b) The experimental units are Facebook users.
c) The explanatory variable is manipulated versus not manipulated. The response variables are percentages of all words that were positive and percentage of all words that were negative.
d) If the participants were informed, they would not have acted authentically. They were likely upset that they were not informed because that violates the requirement of informed consent.

4.36 Texting while driving

a) Is there an association (or relationship) between texting while driving and driver control ability?
b) The explanatory variable is texting. The treatments are typing alone, reading alone, or a combination of both. The response variables are eye movements, stimulus detection, reaction time, collisions, lane positioning, speed and headway. The experimental units are drivers.
c) Combined efforts including legislation, enforcement, blocking technologies, parent modeling, social media, social norms, and education will be required to prevent continued deaths and injuries from texting and driving.

4.37 Pain reduction medication

It is important to use a placebo so that the two treatment groups appear identical to the individuals in the study. It is also important to account for the placebo effect, people who take a placebo tend to respond better than those who take nothing at all. Without a placebo or a control comparison group, there is no way to separate the placebo effect from the actual effect of the medication.

4.38 Pain reduction medication, continued

a) The second design is better for generalizing the results to the entire population. Under the first design, it is impossible to tell whether the results are due to the medication or gender.
b) As long as the recruited individuals are representative of the population, the results can be generalized.

4.39 Pain reduction medication, yet again

The researchers should be blinded to the treatment as well so that they don't intentionally or unintentionally treat the subjects differently according to which treatment group they are in.

4.40 Colds and vitamin C

a) (i) We would have two treatments: vitamin C and a placebo version of vitamin C.
   (ii) Subjects would be randomly assigned to the two treatments using random numbers. They would be asked to take a pill (either vitamin C or placebo) every day for an entire winter. The presence of vitamin C would be the explanatory variable and the presence of the common cold over the next winter would be the response variable.
   (iii) We could make the study double blind by not letting subjects know if the pill they were taking was vitamin C or a placebo. In addition, the people from the research team who contact participants to find out if they had had a cold would not know which treatment the subjects were in.
b) People who regularly take vitamin C might also be more health conscious and have other healthy behaviors. They might exercise, get plenty of sleep, and eat nutritious foods that have all kinds of vitamins – not just vitamin C. We would not be able to isolate vitamin C from these other possible lurking variables.

4.41 Reducing high blood pressure

a) We could design an experiment by recruiting volunteers with a history of blood pressure; these volunteers would be the experimental units. The volunteers could be randomly assigned to one of two treatments: the new drug or the current drug. In this experiment, the explanatory variable would be treatment type and the response variable would be blood pressure after the experimental period.
b) To make the study double-blind, the two drugs would have to look identical so that neither the subjects nor the experimenters who have contact with subjects know what drug a particular subject is taking.

Section 4.4 Other Ways to Conduct Experimental and Nonexperimental Studies

4.42 Student loan debt

This is a stratified sample because the sample sizes are fixed in each of the groups of interest, graduates of four-year public universities and graduates of four-year private universities. The two types of universities are called strata. Simple random samples of 100 graduates are then taken from each type of university.

4.43 Club officers again

a) This sample is drawn by numbering the students from 1 to 5.Then, one-digit numbers are randomly picked, and the students selected are the first female student to have her number picked (1 to 3) and the first male student to have his number picked (4 or 5). Numbers beyond the range of 1 to 5 are ignored, as are duplicates. In addition, once a student has been chosen, we will ignore the numbers of other students of that gender. For example, if 2 is picked, we would then ignore 1 and 3, because the other student must be male.

b) This is not a simple random sample because every sample of size two does not have an equal chance of being selected. For example, samples of two women or two men are prohibited from being selected at all. Moreover, this means that each of the two men has a higher chance of being selected than does any of the three women.

c) Since there are three females, each having an equally likely chance of being selected and one must be selected, the activity coordinator has a 1 in 3 chance of being selected. If the activity coordinator is male, he has a 1 in 2 change of being selected since there are two males.

4.44 Security awareness training

a) A stratified random sample can be used to obtain the desired random sample consisting of $0.25 \times 20 = 5$ employees from production, $0.40 \times 20 = 8$ employees from sales and marketing and $0.35 \times 20 = 7$ employees from new product development.

b) The answer to this problem is based on a random process. This leads to potentially different answers each time it is performed. First, label the employees from 001 to 400 where those numbered 001 to 100 work in production, those numbered 101 to 260 work in sales and marketing and those numbered 261 to 400 work in new product development. We then select five numbers from 01 to 100, eight from 101 to 260 and seven from 261 to 400. As we select random three-digit numbers, we ignore those outside of our range, as well as duplicates. Additionally, once we have selected the correct number from a stratum, we only choose employees from the remaining strata.

4.45 Teaching and learning model

a) Label the 24 schools from 01 to 24. Randomly select schools until roughly 20% of the students (about $0.25 \times 6057 = 1212$ students) are included in the sample. Note that how many schools are included in the sample will vary from one sample to the next since the schools have different numbers of students.

b) Random samples will vary.

c) Random samples will vary.

d) No, it would not be possible to implement a stratified random sample in this case. The only possible strata are the schools themselves and as mentioned previously, the new model must be implemented at the entire school, it cannot be implemented for a random sample of students selected from each school.

4.46 Pelvic girdle pain and sick leave

a) A retrospective study is one in which subjects are asked to report on their sick leaves when they were pregnant.

b) Cases refer to subjects who had PGP, and controls refer to subjects who did not have PGP.

c) (i) 193/221 = 0.87 (proportion of subjects with PGP who had been on sick leave during pregnancy).
(ii) 236/347 = 0.68 (proportion of subjects who did not have PGP who had been on sick leave).

4.47 Smoking and lung cancer

It is possible that people with lung cancer had a different diet than did those without. For example, these people might have eaten out at restaurants quite a bit, thus consuming more fat. The social aspect of eating out might also have made them more likely to smoke. However, it could have been the fat and not the smoking that caused lung cancer.

4.48 Dream job and reality

This study was a prospective study because subjects were followed for 15 years. They reported their dream jobs when they were teenagers, and they were followed to see if they have had their dream jobs 15 years later.

4.49 Baseball under a full moon

a) Yes, in large databases it is possible to uncover many surprising trends.
b) Yes, anticipated results stated in advance are generally more convincing than results that have already occurred.
c) In a prospective study, the researcher can gather the data he/she desires in the manner he/she desires. In a retrospective study, variables that the researcher is interested in may not be available and the data may not have been collected consistently or accurately. For these reasons, a prospective study will likely give more reliable results.

4.50 Are two factors helpful?

a) One factor is crop type, and the other is type of fertilizer. The response variable is the total yield in lbs/acre.
b) If the study did a one-factor analysis of fertilizer, it would have found significant better results for the inorganic fertilizer with microbes, and would have concluded that type of fertilizer makes a significant difference.
c) From the two-factor study, we can learn that the inorganic fertilizer with microbes is significantly efficient for cucumber crop, whereas for okra crop, no significant difference between the two types of fertilizer was shown.

   We would have concluded that the type of fertilizer makes a difference if we had only looked at the type of fertilizer and not on the crop type.

4.51 Growth Mindset

a) The explanatory variable is type of praise. The treatments are praised for effort versus praised for intelligence. The response variable is whether they chose a challenging task on the subsequent task. The experimental units are the study participants.
b) This is a randomized experiment.

4.52 Which teaching method is more efficient?

a) The blocks are the students. This type of block design is called a matched design.
b) If this study was double-blind that would mean that neither students nor experimenters who have contacted them knew whether they were enrolled in the section that make use of the traditional, the hybrid or the online delivery method which could not be realistic in this case.
c) This study could incorporate randomization by randomly assigning subjects to a given order of delivery method for the parts of the course. For example, one subject might be randomly assigned to the online delivery method for the first part of the course, then the traditional method for the second part, then the hybrid method of the third. Another subject might be randomly assigned to start with hybrid, then follows with the traditional, then ends with the online. And so on. This helps to reduce possible bias.

4.53 Effect of partner smoking in smoking cessation study

a) This is not a completely randomized design because the researchers are not randomly assigning subjects to living situation (living with another smoker versus not living with another smoker).
b) The experiment has two blocks: those living with smokers and those not living with smokers.
c) This is a randomized block design because randomization of units to treatments occurs within blocks.

Chapter Problems: Practicing the Basics

4.54 Cell phones

Since the outcome under study is brain cancer, it is more realistic to collect a random sample of subjects who have contracted brain cancer and then look at their past cell phone usage. An experiment would require the researcher to randomly assign subjects to the various treatments (varying amounts of cell phone use) and then observe these subjects long enough for a sufficient number to develop brain cancer. This is not practical and likely not even possible.

4.55 Observational versus experimental study

In an observational study, we observe people in the groups they already are in. For example, we might compare cancer rates among smokers and nonsmokers. We do not assign people to smoke or not to smoke; we observe the outcomes of people who already smoke or do not smoke. In an experiment, we actually assign people to the groups of interest. Although it would be unethical, we could turn the above observational study into an experiment by assigning people either to smoke or not to smoke. We would not allow them to make this choice. (Of course, even if we ignored ethics and did this, our subjects might ignore our instructions!) The major weakness of an observational study is that we cannot control (such as by balancing through randomization) other possible factors that might influence the outcome variable. For example, in the smoking study, it could be that smokers also drink more, and drinking causes cancer. With the experiment, we randomly assign people to smoke or not to smoke; thus, we can assume that these groups are similar on a range of variables, including drinking. If the smokers still have higher cancer rates than the nonsmokers, we can assume it's because of smoking, and not because of other associated variables such as drinking.

4.56 Unethical experimentation

Examples will vary. We might be interested in whether combat experience leads to higher rates of anxiety disorders; however, we cannot randomly assign people to go off to war. We could, however, conduct an observational study comparing rates of anxiety disorders among combat veterans and among a similar sample of non-veterans.

4.57 Spinal fluid proteins and Alzheimer's

a) The explanatory variables are whether the individual has the two types of proteins in their spinal fluid and the state of the individual's memory at the time the spinal fluid was analyzed: normal memory, memory problems or Alzheimer's disease. The response variable is whether or not the individual developed Alzheimer's within the next five years.

b) This is a non-experimental study because individuals are not assigned to a specific memory classification.

c) It would not be practically possible to design this study as an experiment because one cannot assign individuals to a specific memory category.

4.58 Fear of drowning

You should give more weight to the study than to the story, which is just anecdotal evidence. Something can be true on an average, and yet there can still be exceptions, such as the teenager taking the selfie in the middle of the sea you have read about. The story of one person is anecdotal and not as strong evidence as a carefully conducted study with a much larger sample size.

4.59 NCAA men's basketball poll

a) Indiana: $\left(1/\sqrt{n}\right)\times 100\% = \left(1/\sqrt{3300}\right)\times 100\% = 1.7\%$

Wisconsin: $\left(1/\sqrt{n}\right)\times 100\% = \left(1/\sqrt{5600}\right)\times 100\% = 1.3\%$

b) The percentages varied so drastically because of home team support. The residents of Indiana and Wisconsin wanted their teams to win and were willing to overlook their chances of winning in predicting the outcome.

c) One type of potential bias is from undercoverage, the only possible respondents are visitors to the website espn.com, not all residents of the state. Another is from this being a volunteer sample, visitors to the site chose whether or not to respond to the website's survey.

4.60 Sampling your city's citizens

a) This would be an example of bias because some parts of population are favored over others. Moreover, restricting our range in this way would not allow us to determine an association between these variables.

b) There are different ways to select a sample that would yield useful information. Here are two samples. We could select a simple random sample. Or we could get a list of all citizens in the city, number them (e.g., 0001 to 3000 for a city of three thousand citizens), then select 50 random five-digit numbers in this range. Alternately, if we wanted to be sure you had equal numbers of males and females, you could stratify your sample, dividing the citizens into these categories, and selecting 25 citizens from each gender.

4.61 Beware of online polls

Respondents to such an internet poll might be those who feel most strongly about this topic. For example, iPhones and iPads owners who are concerned about their privacy might have felt more compelled to respond than non-iPhones and iPads owners. The lurking variable might be ownership of iPhones and iPads. Owners of iPhones and iPads might be mobilized to cast votes with greater probability than others.

4.62 Comparing female and male students

a) The students would be numbered from 0001 to 3500. Then we would choose random four-digit numbers, ignoring those outside the range of 0001 to 3500, and ignoring duplicates. We would select the first three students whose numbers matched these criteria.

b) No; every sample is not equally likely; any possible sample with more than 40 males or fewer than 40 males has probability 0 of being chosen.

c) This would be a stratified random sample. It offers the advantage of having the same numbers of men and women in the study, which would be unlikely if the population had a small proportion of one of these, and this is useful for making comparisons.

4.63 Second job for MPs

a) Because this is a volunteer sample, there is a potential for sampling bias, both because the sample is not randomly selected (those who responded might have been those who felt the most strongly) and because of undercoverage (anyone without Internet access would not have been able to participate). There is also potential for response bias because the statements are leading.

b) If the sample is biased due to undercoverage and lack of random sampling, it does not matter how big the sample. It is always better to have a small random sample than a large volunteer sample.

4.64 Obesity in metro areas

a) No, we are not able to conclude that obesity causes a higher incidence of these conditions because this is an observational study, not an experiment.

b) Answers will vary. Some possibilities are education level and income.

4.65 Voluntary sports polls

a) No, it was done by voluntary response.

b) They could send out a survey to all sports fans (perhaps using a cluster sampling method by choosing to sample from each sports stadium).

4.66 Video games mindless?

a) The explanatory variable is history of playing video games, and the response variable is visual skills.

b) This was an observational study because the men were not randomly assigned to treatment (played video games versus hadn't played); those who already were in these groups were observed.

c) One possible lurking variable is reaction time. Excellent reaction times might make it easier, and therefore more fun, to play video games, leading young men to be more likely to play. Excellent reaction times also might lead young men to perform better on tasks measuring visual skills. These young men might have performed well on tasks measuring visual skills regardless of whether they played video games.

4.67 Physicians' health study

a) This was (i) an experiment, and it was (ii) prospective.

b) The response variable was presence/absence of myocardial infarction and the explanatory variable was treatment group (aspirin or placebo).

4.68 Aspirin prevents heart attacks?

a) The response variable was whether they had a heart attack; the explanatory variable was treatment group (aspirin or placebo).

b) This is an experiment because physicians were randomly assigned to treatment – either aspirin or placebo.

c) Because the experiment is randomized, we can assume that the groups are fairly balanced with respect to exercise. Each group would have some physicians with low exercise and some with high. On average, they'd be similar.

4.69 Exercise and heart attacks

a) 72.5% of physicians in the aspirin group exercised vigorously, and 72.0% of physicians in the placebo group exercised vigorously. These percentages are very similar.

b) The percentages of physicians in the two groups who exercised vigorously are very similar. It does seem that the randomization process did a good job in achieving balanced treatment groups in terms of exercise. Heart attack response between the two groups should not be systematically influenced in one direction or the other due to exercise.

4.70 Smoking and heart attacks

a) Among those who were in the aspirin group, 49.3% never smoked, 39.7% smoked in the past, and 11.0% are current smokers. Among those in the placebo group, 49.8% never smoked, 39.1% smoked in the past, and 11.1% currently smoke. These proportions are very similar.

b) It does seem that the randomization process did a good job in achieving balanced treatment groups in terms of smoking status. Because there are similar proportions of physicians in both groups who report that they have never smoked, used to smoke, or currently smoke, this variable is not likely to be responsible for any differences in heart attack rates. The heart attack response between the two groups should not be systematically influenced in one direction or the other due to smoking status of the physicians.

4.71 Aspirin, beta-carotene, and heart attacks

Beta-carotene (Factor 2)	Aspirin (Factor 1): Yes	Aspirin (Factor 1): No
Yes	Treatment 1 Aspirin Beta-carotene	Treatment 2 No Aspirin (Placebo) Beta-carotene
No	Treatment 3 Aspirin No beta-carotene (Placebo)	Treatment 4 No aspirin (Placebo) No beta-carotene (Placebo)

4.72 Bupropion and nicotine patch study results

a) Nicotine patch only: $\left(1/\sqrt{244}\right)\times 100\% = 6.4$ percentage points

Bupropion only: $\left(1/\sqrt{244}\right)\times 100\% = 6.4$ percentage points

Nicotine patch with bupropion: $\left(1/\sqrt{245}\right)\times 100\% = 6.4$ percentage points

Placebo only: $\left(1/\sqrt{160}\right)\times 100\% = 7.9$ percentage points

It is believable that the true abstinence percentage falls anywhere within the range indicated by the margin of error. For example, the range for the nicotine patch only is 16.4% – 6.4% = 10% to 16.4% + 6.4% = 22.8%. These are all believable values for the abstinence percentage of those using the nicotine patch only.

b) Yes, it does seem as if the treatments bupropion only and Placebo only are different. The margin of error for bupropion only indicates that the low end of believable values is 30.3% – 6.4% = 23.9%, whereas the margin of error for Placebo only indicates that the high end of believable values is 15.6% + 7.9 = 23.5%. Because there's no overlap, we can conclude that it's likely that these two percentages are significantly different from one another.

c) No, it does not seem as if the treatments bupropion only and Nicotine patch with bupropion are significantly different. There is substantial overlap between the ranges indicated by the margins of error. The range for bupropion only extends from 23.9% to 36.7%, and the range for Nicotine patch with bupropion extends from 29.1% to 41.9%. Because 29.1% through 36.7% are believable values for both treatments, we cannot conclude that there are different abstinence percentages for these two groups.

d) Using the results of (a) – (c), the results of the study suggest that two of the treatments, bupropion only and Nicotine patch with bupropion, led to higher abstinence percentages than did either of the other two treatments, Nicotine only or Placebo only. However, there was not a statistically significant difference between bupropion only and Nicotine patch with bupropion. One possible recommendation may be to use bupropion only as an aid for quitting smoking.

4.73 Prefer M&Ms or Smarties?

a) (i) If you wanted to use a completely randomized design, you could randomly assign subjects to one of the two treatments, M&Ms or Smarties, and have them rate it on a scale such as 0 to 10 with higher numbers being a better rating. Neither the subject nor the experimenter would know which candy the subject was tasting. (For example, an experimenter with no contact with the subjects could put both types of candies into identical bags; a code could be used to later identify the type of candy.)

(ii) If one wanted to use a matched-pairs design, we could have all subjects participate in both the treatments. We would randomly assign each subject to taste either M&Ms or Smarties first, then to taste the other one, and then to indicate which they prefer. As in the completely randomized design, neither the subject nor the experimenter administering the candy would know which type of candy the subject was tasting.

b) There are advantages to each. The completely randomized design eliminates the possibility that tasting one candy would alter one's preference for the second candy. The matched-pairs design decreases the possible effects of lurking variables because the two groups are made up of the same people.

4.74 Comparing gas brands

a) The response variable is the gas mileage. The explanatory variable is the brand of gas. Its treatments are Brand A (the name brand) and Brand B (the independent brand).

b) In a completely randomized design, 10 cars would be randomly assigned to Brand A and 10 cars to Brand B.

c) In a matched-pairs design, each car would be a block. It would first use gas from one brand, and then from the other.

d) A matched-pairs design would reduce the effects of possible lurking variables because the two groups would be identical. With a completely randomized design, it is possible that, just by chance, one group of cars gets better gas mileage to begin with than does the other group.

4.75 Samples not equally likely in a cluster sample?

With a simple random sample, every possible *sample* of a given size has an equal chance of being selected. With a cluster random sample, even with equally-sized clusters, there are many samples that have *no* chance of being selected. As just one example, a sample that has one subject from each cluster would never be selected in a cluster random sample because cluster random sampling, by definition, includes *all* the subjects in the chosen clusters. For example, the cluster sample will not contain subjects from every cluster.

4.76 Nursing homes

a) The nursing homes are clusters.

b) The sample is not a simple random sample because each possible sample does not have an equal chance. For instance, there is no chance of a sample in which just one person from a particular nursing home is sampled.

4.77 Multistage health survey

The first stage is the division of the U.S. into four regions – these are strata, and so the researcher is using stratification to ensure she has the same number of subjects from each region. She then takes a simple random sample of ten schools in each region. At this level, she is using cluster sampling, because she is identifying all possible clusters (schools), and randomly selecting a given number of them in each region. She then randomly samples three classrooms in each school. This also is cluster sampling because she is identifying all possible clusters (classrooms this time), and randomly selecting a given number of them in each selected school. Finally, she interviews all students in those classrooms.

4.78 Hazing

This is cluster random sampling. The colleges are the clusters.

4.79 Marijuana and schizophrenia

It seems that there are lurking variables that are responsible for at least some of the association. For example, it is possible that individuals who are genetically susceptible to schizophrenia are also predisposed to liking marijuana. Education level and socioeconomic status are also possible lurking variables.

4.80 Family disruption and age at menarche

a) This was a retrospective case-control study because it recruited patients based on age at menarche, and then investigated their past in terms of family disruption/father absence prior to the menarche event.

b) Subjects who are biological sister pairs would have the same or similar values of many possible lurking variables.

Chapter Problems: Concepts and Investigations

💻4.81 Cell phone use

The answers to these questions will be different for each student, depending on the study that each student locates.

💻4.82 Read a medical journal

The answers to these questions will be different for each student, depending on the study that each student locates.

💻4.83 Internet poll

Regardless of the study found, the results should not be trusted due to the volunteer nature of the sample.

4.84 Search for an observational study

The answers to these questions will be different for each student, depending on the study that each student locates.

4.85 Search for an experimental study

The answers to these questions will be different for each student, depending on the study that each student locates.

4.86 Judging sampling design

a) This is a volunteer survey, and so likely suffers from sampling bias. It is quite possible, for example, that those with the most need for these social programs have the least access to this Internet poll. This type of sampling bias is called undercoverage because some groups in the population are not in the sampling frame.
b) Again, this is a volunteer sample. This is not a random sample. Those who are writing to the congresswomen are probably those who have the strongest opinions about this issue.
c) This is a biased sample. Physical and social science majors who have chosen to take a course in Comparative Human Sexuality quite possibly already are more similar in terms of sexual attitudes than are physical and social science majors who have not chosen to take this course.
d) This study suffers from nonresponse bias. It is quite possible that the very large percentage of people who did not respond are different in some important way from the smaller proportion who did respond.

4.87 More poor sampling designs

a) The principal is attempting to use cluster sampling by listing all of her clusters (first-period classes), and taking a random sample of clusters. However, her sample includes only one cluster. She would need to choose several clusters in order to have something resembling a representative sample.
b) Values might be higher than usual because the days sampled are at the start of the weekend. Sampling just Fridays is an example of sampling bias. She should take a simple random sample of all days in the past year.

4.88 Age for legal alcohol

a) "Do you think it should be legal for people to drink at age 18 given that they can get married, go to war, drive a car, and buy cigarettes by this age?" This is biased because it pushes respondents to say yes.
b) "Do you think that it should be legal or illegal for people to drink at age 18?"

4.89 Quota sampling

This is not a random sampling method. People who approach the street corner are interviewed as they arrive (and as they agree to the interview!). Although researchers strive to obtain data from people from a number of backgrounds, the people within these backgrounds (e.g., Hispanic) who are surveyed on the street corner may not be representative of the general population of that kind of person. Although the quota leads to a diversity of people being surveyed, the choice of a given street corner likely constitutes sampling bias.

4.90 Smoking and heart attacks

a) This is an observational study. No subject was assigned to a treatment.
b) There could have been numerous lurking variables that might explain this association. For example, perhaps the first six months included the summer and perhaps the population size in that part of Montana is much higher in the summer.

4.91 Issues in clinical trials

a) Randomization is necessary because subjects would choose the treatment in which they have the most faith. Such a study would be a measure of how well a treatment works if patients believe in it, rather than how much a treatment works independent of subjects' beliefs about its efficacy.
b) Patients might be reluctant to be randomly assigned to one of the treatments because they might perceive it as inferior to another treatment. In this case, patients might perceive (even in the absence of the data that this study is trying to collect) that the new treatment will be an improvement, and might be reluctant to participate in the study without the guarantee that they can get that treatment.
c) If the researcher thinks that the new treatment is better than the current standard, he or she might be reluctant to proceed because he or she might feel that <u>all</u> patients should get the new treatment, and not just those randomly assigned to it.

4.92 Compare smokers with nonsmokers?

This study has equal numbers of people with and without lung cancer. The rates of lung cancer are not nearly so high in the general population; these samples were chosen specifically because of their lung cancer status. Thus, in this sample, we will find far more smokers <u>and</u> nonsmokers with lung cancer than in the general population, because we have explicitly chosen a sample with such a high rate of lung cancer.

4.93 Is a vaccine effective?

Because the disease is so rare, it's very unlikely that the 200 people randomly chosen to be in this study would have the disease, whether or not they get the vaccine. It would be more practical to find a certain number of people who already have the rare disease (the cases). We would compare the proportion of these people who had received the vaccine to the proportion in a group of controls who did not have the disease.

4.94 Distinguish helping and hindering among infants

a) If the videos are always shown in the same order, it's possible that the infant grows bored before the second video and simply chooses the toy he/she recognizes from the first video.

b) Randomizing the order in which the videos are shown will minimize this type of response bias lending more credibility to the results.

💻4.95 Distinguish helping and hindering among infants, continued

Answers will vary. Number the infants 01 to 16. Using the table of random digits, select distinct pairs of two digit numbers between 01 and 16 until eight have been chosen. To use a random number generator, select the minimum value as 1, the maximum value as 16, and the number to select to eight. These eight infants are assigned to watch the video with the helpful figure first. The remaining eight infants are assigned to watch the video with the hindering figure first.

💻4.96 Distinguish helping and hindering among infants, continued

Answers will vary but are most likely to be either 0 or 1. The results should provide convincing evidence that infants actually tend to exhibit a preference.

4.97 Multiple choice: What's a simple random sample?

The best answer is (b).

4.98 Multiple choice: Be skeptical of medical studies?

The best answer is (b).

4.99 Multiple choice: Opinion and question wording

The best answer is (a).

4.100 Multiple choice: Campaign funding

The best answer is (b).

4.101 Multiple choice: Emotional health survey

The best answer is (d).

4.102 Multiple choice: Sexual harassment

The best answer is (a).

4.103 Multiple choice: Effect of response categories

The best answer is (b).

♦♦4.104 Systematic sampling

a) Although every subject is equally likely to be chosen – at least before the first subject is chosen – every possible sample of 100 is not equally likely. We would never, for example, have a sample that included subjects whose names were next to each other on the population list.

b) The company would determine the first item using a randomly selected two-digit number between 01 and 50; the item that coincided with that number would be checked, and then every 50th item thereafter also would be checked.

♦♦💻4.105 Complex multistage GSS sample

(Note: Because there are many aspects of this study, and sampling changes over the years of the study, there are several possible responses for each part of this exercise.)

a) In some of the research, clustering was used. For example, the researchers randomly selected among all possible Standard Metropolitan Statistical Areas (SMSAs) or non-metropolitan counties.

b) Much of the research used stratification. For example, within the clusters mentioned above, the researchers stratified participants based on region, age, and race.

c) Simple random sampling also was used at times. The GSS website notes: "The full-probability GSS samples used since 1975 are designed to give each household an equal probability of inclusion in the sample." Simple random sampling means that each subject (in this case, household), has an equal chance of being in the study.

♦♦4.106 Mean family size

Consider the hint. The mean population family size is 6. If we choose families, the possible values are 2 and 10, each equally likely, and the sampling is not biased. If we choose individuals, the value 2 has probability 2/12 and the value 10 has probability 10/12. We are likely to overestimate the mean size.

♦♦4.107 Capture–recapture

a) M is the 50 deer who were tagged initially, n is the 125 deer who were captured several weeks later, and R is 12, the number of tagged deer who showed up in the second sample.

b) If one assumes that the sample proportion of tagged deer equals the population proportion of tagged deer, one can easily calculate the estimated population size of deer. We know the proportion of tagged deer in the sample is the number of tagged deer in the second sample (R), divided by the total number of deer in the second sample (n). We also know the total number of tagged deer in the population (M). That leaves only one variable to solve for, the estimated number of deer in the population (N).

c) $\frac{R}{n} = \frac{M}{N} \Rightarrow \frac{12}{125} = \frac{50}{N} \Rightarrow N = \frac{125 \cdot 50}{12} = 520.8$

d)

		In first sample (the census itself)?		
		Yes (returned form)	**No (did not return form)**	**Total**
In second sample (PES)?	**Yes**	***R***		***n***
	No			
	Total	***M***		***N***

Chapter Problems: Student Activities

4.108 Munchie capture–recapture

If the estimate is not close, one factor that could be responsible is the fairly small sample size. This could be a problem in real-life applications as well, particularly if the animal species of interest were endangered.

💻4.109 Activity: Sampling the states

Answers will vary.