

Chapter 3 Solutions

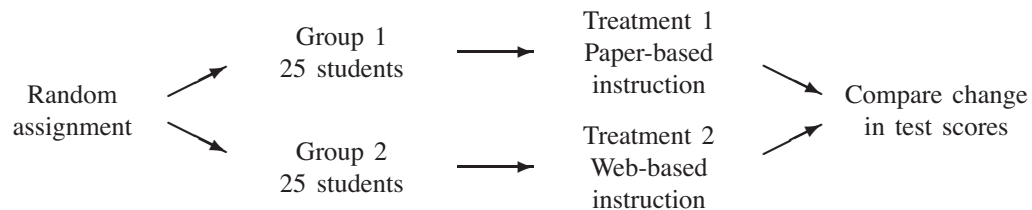
- 3.1.** Any group of friends is unlikely to include a representative cross section of all students.
- 3.2.** Political speeches provide a good source of examples.
- 3.3.** A hard-core runner (and her friends) are not representative of all young people.
- 3.4.** The performance of one car is anecdotal evidence—a “haphazardly selected individual case.” People tend to remember—and re-tell—stories about extraordinary performance, while cases of average or below-average reliability are typically forgotten.
- 3.5.** For example, who owns the Web site? Do they have data to back up this statement, and if so, what was the source of that data?
- 3.7.** This is an experiment: Each subject is assigned to a treatment group (presumably at random, although the description does not tell us this is the case). The explanatory variable is the drug received, the response variables are adverse events, as well as some reaction. (The nature of that reaction is not specified in the exercise, but they apparently collected some information to indicate which subjects had an “effective response” to the vaccine.)
- 3.8. (a)** No treatment is imposed on the subjects (children); they (or their parents) choose how much TV they watch. The explanatory variable is hours watching TV, and the response variable is “later aggressive behavior.” **(b)** An adolescent who watches a lot of television probably is more likely to spend less time doing homework, playing sports, or having social interactions with peers. He or she may also have less contact with or guidance from his/her parents.
- 3.9.** This is an experiment: Each subject is (presumably randomly) assigned to a group, each with its own treatment (computer animation or reading the textbook). The explanatory variable is the teaching method, and the response variable is the change in each student’s test score.
- 3.10.** This is an experiment, assuming the order of treatment given to each subject was randomly determined. The explanatory variable is the form of the apple (whole, or juice), and the response variable is how full the subjects felt.
Note: *This is a matched pairs experiment, described on page 181 of the text.*
- 3.11.** The experimental units are food samples, the treatment is exposure to different levels of radiation, and the response variable is the amount of lipid oxidation. Note that in a study with only one factor—like this one—the treatments and factor levels are essentially the same thing: The factor is varying radiation exposure, with nine levels.
It is hard to say how much this will generalize; it seems likely that different lipids react to radiation differently.

3.12. This is an experiment because the experimental units (students) are randomly assigned to a treatment group. Note that in a study with only one factor—like this one—the treatments and factor levels are essentially the same thing: There are two treatments/levels of the factor “instruction method.” The response variable is the change in score on the standardized test.

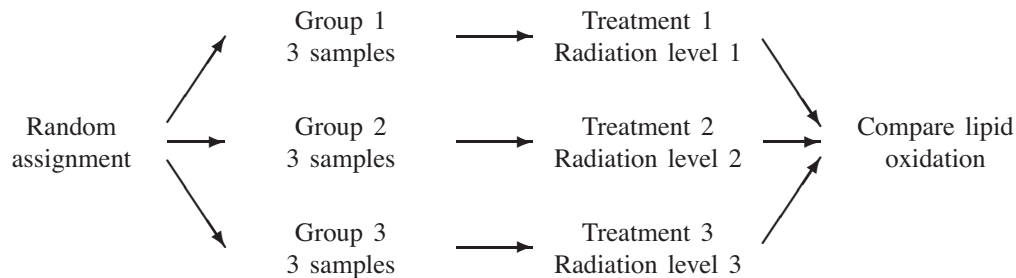
The results of this experiment should generalize to other classes (on the same topic) taught by the same instructor, but might not apply to other subject matter, or to classes taught by other instructors.

3.13. Those who volunteer to use the software may be better students (or worse). Even if we cannot decide the direction of the bias (better or worse), the lack of random allocation means that the conclusions we can draw from this study are limited at best.

3.14.



3.15. Because there are nine levels, this diagram is rather large (and repetitive), so only the top three branches are shown.

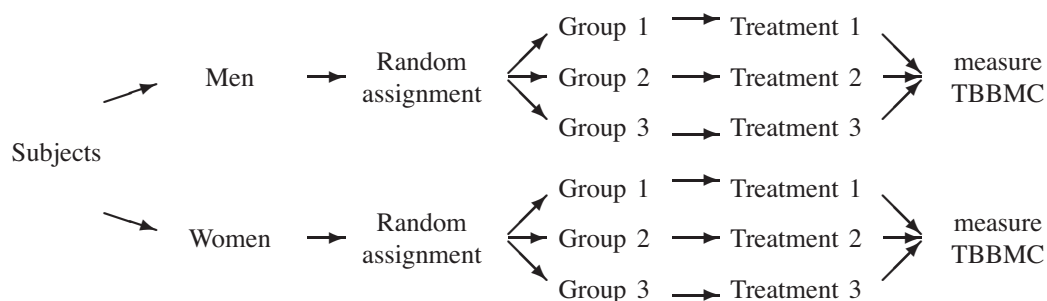


3.16. The results will depend on the software used.

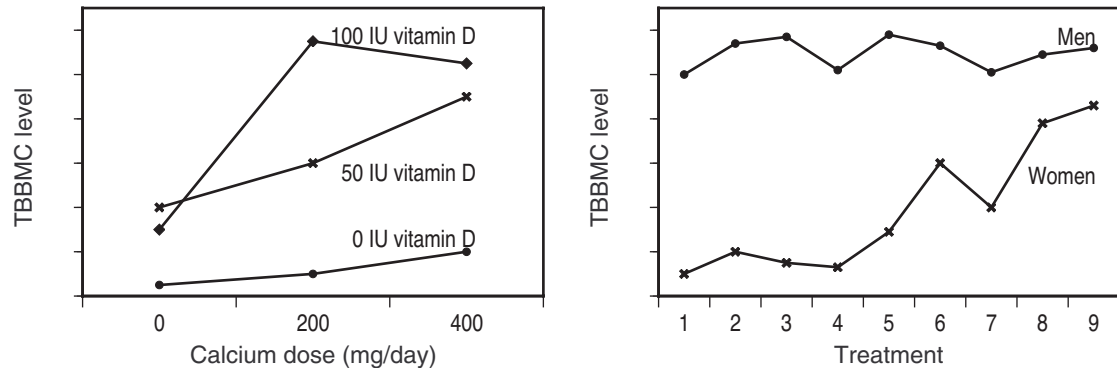
3.17. (a) Shopping patterns may differ on Friday and Saturday, which would make it hard to determine the true effect of each promotion. (That is, the effect of the promotion would be confounded with the effect of the day.) To correct this, we could offer one promotion on a Friday, and the other on the following Friday. (Or, we could do as described in the exercise, and then on the next weekend, swap the order of the offers.) **(b)** Responses may vary in different states. To control for this, we could launch both campaigns in (separate) parts of the same state or states. **(c)** A control is needed for comparison; if we simply compare this year’s yield to last year’s yield, we will not know how much of the difference can be attributed to changes in the economy. We should compare the new strategy’s yield with another investment portfolio using the old strategy.

Note: For part (c), this comparison might be done without actually buying or selling anything; we could simply compute how much money would have been made if we had followed the new strategy; that is, we keep a “virtual portfolio.” This assumes that our buying and selling is on a small enough scale that it does not affect market prices.

- 3.18. (a)** Assigning subjects by gender is not random. It would be better to treat gender as a blocking variable, assigning five men and five women to each treatment. **(b)** This randomization will not necessarily divide the subjects into two groups of five. (Note that it *would* be a valid randomization to use this method until one group had four subjects, and then assign any remaining subjects to the other group.) **(c)** The 10 rats in a batch might be similar to one another in some way. For example, they might be siblings, or they might have been exposed to unusual conditions during shipping. (The safest approach in this situation would be to treat each batch as a block, and randomly assign two or three rats from each batch to each treatment.)
- 3.19.** The experiment can be single-blind (those evaluating the exams should not know which teaching approach was used), but not double-blind, because the students will know which treatment (teaching method) was assigned to them.
- 3.20.** For example, we might block by gender, by year in school, or by housing type (dorm/off-campus/Greek).
- 3.21.** For example, new employees should be randomly assigned to either the current program or the new one. There are many possible choices for outcome variables, such as performance on a test of the information covered in the program, or a satisfaction survey or other evaluation of the program by those who went through it.
- 3.22.** Subjects—perhaps recruited from people suffering from chronic pain, or those recovering from surgery or an injury—should be randomly assigned to be treated with magnets, or a placebo (an object similar to the magnets, except that it is not magnetic. Students should address some of the practical difficulties of such an experiment, such as: How does one measure pain relief? How can we prevent subjects from determining whether they are being treated with a magnet? (For the latter question, we might apply the treatments in a controlled setting, making sure that there is nothing metal with which the subjects could test their treatment object.)
- 3.23. (a)** The factors are calcium dose, and vitamin D dose. There are 9 treatments (each calcium/vitamin D combination). **(b)** Assign 20 students to each group, with 10 of each gender. The complete diagram (including the blocking step) would have a total of 18 branches, below is a portion of that diagram, showing only three of the nine branches for each gender. **(c)** Randomization results will vary.



3.24. Students may need guidance as to how to illustrate these interactions. Figure 3.7 shows one such illustration (as part of Exercise 3.40). Shown below are two possible illustrations, based on made-up data (note there is no scale on the vertical axis). In the first, we see that the effect of vitamin D on TBBMC depends on the calcium dose. In the second, we see little variation in men's TBBMC across all nine treatments, while women's TBBMC appears to depend on the treatment group. (In particular, women's TBBMC is greatest for treatment 9, with the highest doses of both calcium and vitamin D.)



3.25. (a) For example, flip a coin for each customer to choose which variety (s)he will taste. To evaluate preferences, we would need to design some scale for customers to rate the coffee they tasted, and then compare the ratings. **(b)** For example, flip a coin for each customer to choose which variety (s)he will taste *first*. Ask which of the two coffees (s)he preferred. **(c)** The matched-pairs version described in part (b) is the stronger design; if each customer tastes both varieties, we only need to ask which was preferred. In part (a), we might ask customers to rate the coffee they tasted on a scale of (say) 1 to 10, but such ratings can be wildly variable (one person's "5" might be another person's "8"), which makes comparison of the two varieties more difficult.

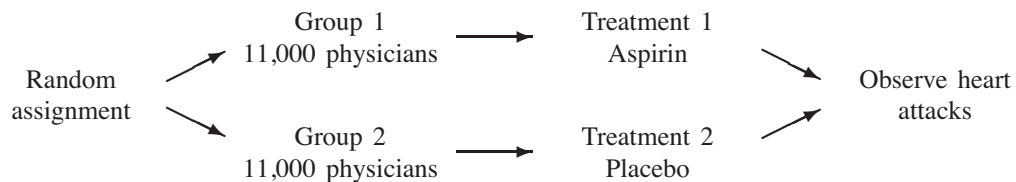
3.26. An experiment would be nearly impossible in this case, unless the publishers of *Sports Illustrated* would allow you to choose the cover photo for several issues. (The ideal design would involve taking a collection of currently-successful teams or individuals, and randomly assigning half to be on the cover, then observing their fortunes over the next few months.)

3.27. Experimental units: pine tree seedlings. Factor: amount of light. Treatments (two): full light, or shaded to 5% of normal. Response variable: dry weight at end of study.

3.28. Experimental units: Middle schools. Factors: Physical activity program, and nutrition program. Treatments (four): Activity intervention, nutrition intervention, both interventions, and neither intervention. Response variables: Physical activity and lunchtime consumption of fat.

3.29. Subjects: adults (or registered voters) from selected households. Factors: level of identification, and offer of survey results. Treatments (six): interviewer's name with results, interviewer's name without results, university name with results, university name without results, both names with results, both names without results. Response variable: whether or not the interview is completed.

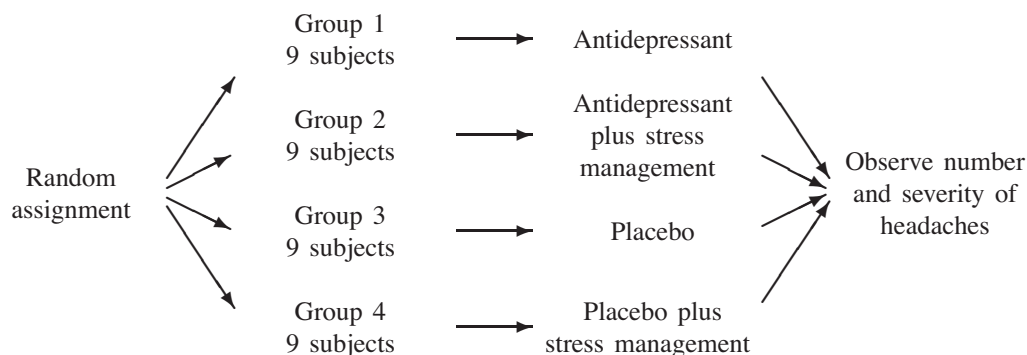
- 3.30. (a)** The subjects are the physicians, the factor is medication (aspirin or placebo), and the response variable is health, specifically whether the subjects have heart attacks. **(b)** Below. **(c)** The difference in the number of heart attacks between the two groups was so great that it would rarely occur by chance if aspirin had no effect.



- 3.31.** Assign nine subjects to treatment 1, then nine more to treatment 2, etc. A diagram is on the next page; if we assign labels 01 through 36, then line 151 gives:

Group 1		Group 2		Group 3	
03 Bezawada	12 Hatfield	32 Tyner	27 Rau	05 Cheng	13 Hua
22 Mi	11 Guha	30 Tang	20 Martin	16 Leaf	25 Park
29 Shu	31 Towers	09 Daye	06 Chronopoulou	28 Saygin	19 Lu
26 Paul	21 Mehta	23 Nolan	33 Vassilev	10 Engelbrecht	04 Cetin
01 Anderson		07 Codrington		18 Lipka	

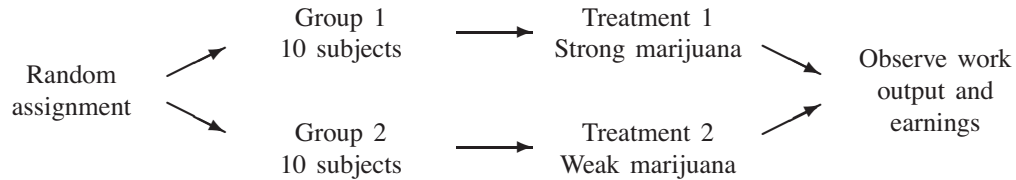
The other nine subjects are in Group 4. The names listed here assume that labels are assigned alphabetically (across the rows). See note on page 50 about using Table B.



3.32. (a) A diagram is shown below. **(b)** Label the subjects from 01 through 20. From line 101, we choose:

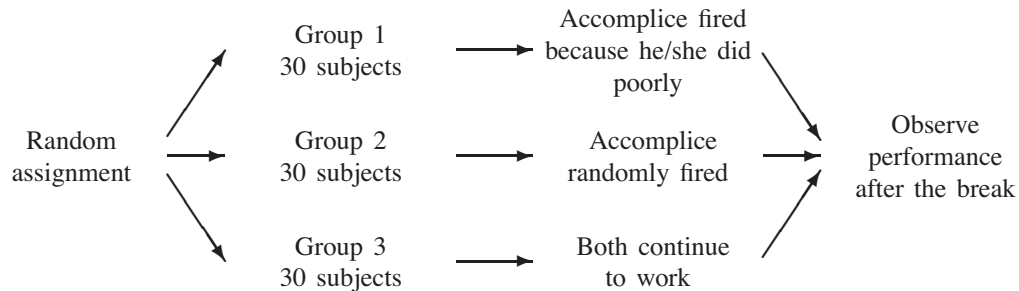
19, 05, 13, 17, 09, 07, 02, 01, 18, and 14

Assuming that labels are assigned alphabetically, that is Wayman, Cunningham, Mitchell, Seele, Knapp, Fein, Brifcani, Becker, Truong, and Ponder for one group, and the rest for the other. See note on page 50 about using Table B.



3.33. Students might envision different treatments; one possibility is to have some volunteers go through a training session, while others are given a written set of instructions, or watch a video. For the response variable(s), we need some measure of training effectiveness; perhaps we could have the volunteers analyze a sample of lake water and compare their results to some standard.

3.34. (a) Diagram below. **(b)** Using line 153 from Table B, the first four subjects are 07, 88, 65, and 68. See note on page 50 about using Table B.



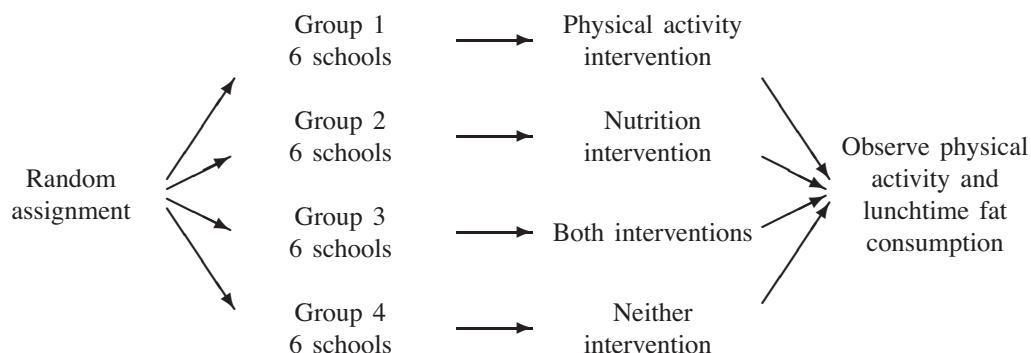
3.35. Diagram below. Starting at line 160, we choose:

16, 21, 06, 12, 02, 04 for Group 1

14, 15, 23, 11, 09, 03 for Group 2

07, 24, 17, 22, 01, 13 for Group 3

The rest are assigned to Group 4. See note on page 50 about using Table B.



3.36. (a) The table below shows the 16 treatments—four levels for each of the two factors.

(b) A diagram is not shown here (with 16 treatments, it is quite large). Six subjects are randomly assigned to each treatment; they read the ad for that treatment, and we record their attractiveness ratings for the ad. Using line 111, the first six subjects are 81, 48, 66, 94, 87, and 60.

		Factor B			
		Fraction of shoes on sale			
		25%	50%	75%	100%
Factor A Discount level	20%	1	2	3	4
	40%	5	6	7	8
	60%	9	10	11	12
	80%	13	14	15	16

3.37. (a) Population = 1 to **150**, Select a sample of size **25**, click **Reset** and **Sample**.

(b) Without resetting, click **Sample** again. **(c)** Click **Sample** three more times.

3.38. Population = 1 to **40**, Select a sample of size **20**, then click **Reset** and **Sample**.

3.39. Design (a) is an experiment. Because the treatment is randomly assigned, the effect of other habits would be “diluted” because they would be more-or-less equally split between the two groups. Therefore, any difference in colon health between the two groups could be attributed to the treatment (bee pollen or not).

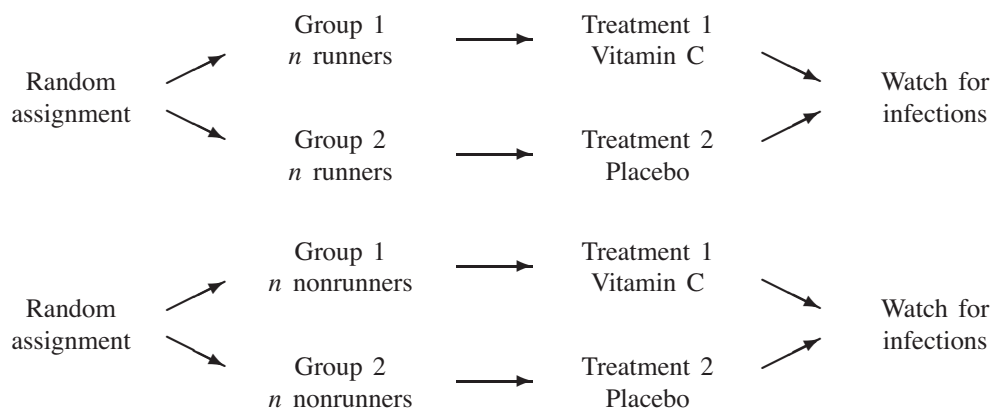
Design (b) is an observational study. It is flawed because the women observed chose whether or not to take bee pollen; one might reasonably expect that people who choose to take bee pollen have other dietary or health habits that would differ from those who do not.

- 3.40.** For a range of discounts, the attractiveness of the sale decreases slightly as the percentage of goods on sale increases. (The decrease is so small that it might not be significant.) With precise discounts, on the other hand, mean attractiveness increases with the percentage on sale. Range discounts are more attractive when only 25% of goods are marked down, while the precise discount is more attractive if 75% or 100% of goods are discounted.
- 3.41.** As described, there are two factors: ZIP code (three levels: none, five-digit, nine-digit) and the day on which the letter is mailed (three levels: Monday, Thursday, or Saturday) for a total of nine treatments. To control lurking variables, aside from mailing all letters to the same address, all letters should be the same size, and either printed in the same handwriting or typed. The design should also specify how many letters will be in each treatment group. Also, the letters should be sent randomly over many weeks.
- 3.42.** Results will vary, but probability computations reveal that more than 97% of samples will have 7 to 13 fast-reacting subjects (and 99.6% of samples have 8 to 14 fast-reacting subjects). Additionally, if students average their 10 samples, nearly all students (more than 99%) should find that the average number of fast-reacting subjects is between 8.5 and 11.5.
Note: X , the number of fast-reacting subjects in the sample, has a hypergeometric distribution with parameters $N = 40$, $r = 20$, $n = 20$, so that $P(7 \leq X \leq 13) \doteq 0.974$. The theoretical average number of fast-reacting subjects is 10.
- 3.43.** Each player will be put through the sequence (100 yards, four times) twice—once with oxygen and once without, and we will observe the difference in their times on the final run. (If oxygen speeds recovery, we would expect that the oxygen-boosted time will be lower.) Randomly assign half of the players to use oxygen on the first trial, while the rest use it on the second trial. Trials should be on different days to allow ample time for full recovery.
 If we label the players 01 through 20 and begin on line 140, we choose 12, 13, 04, 18, 19, 16, 02, 08, 17, 10 to be in the oxygen-first group. See note on page 50 about using Table B.
- 3.44.** The sketches requested in the problem are not shown here; random assignments will vary among students. **(a)** Label the circles 1 to 6, then randomly select three (using Table B, or simply by rolling a die) to receive the extra CO_2 . Observe the growth in all six regions, and compare the mean growth within the three treated circles with the mean growth in the other three (control) circles. **(b)** Select pairs of circles in each of three different areas of the forest. For each pair, randomly select one circle to receive the extra CO_2 (using Table B or by flipping a coin). For each pair, compute the difference in growth (treated minus control).
- 3.45. (a)** Randomly assign half the girls to get high-calcium punch; the other half will get low-calcium punch. The response variable is not clearly described in this exercise; the best we can say is “observe how the calcium is processed.” **(b)** Randomly select half of the girls to receive high-calcium punch first (and low-calcium punch later), while the other half gets low-calcium punch first (followed by high-calcium punch). For each subject, compute the difference in the response variable for each level. This is a better design because it deals with person-to-person variation; the differences in responses for 60 individuals gives more precise results than the difference in the average responses for two groups of 30

subjects. (c) The first five subjects are 38, 44, 18, 33, and 46. In the CR design, the first group receives high-calcium punch all summer; in the matched pairs design, they receive high-calcium punch for the first part of the summer, and then low-calcium punch in the second half.

3.46. (a) False. Such regularity holds only in the long run. If it were true, you could look at the first 39 digits and know whether or not the 40th was a 0. (b) True. All pairs of digits (there are 100, from 00 to 99) are equally likely. (c) False. Four random digits have chance $1/10000$ to be 0000, so this sequence will occasionally occur. 0000 is no more or less random than 1234 or 2718 or any other four-digit sequence.

3.47. (a) This is a block design. (b) The diagram might be similar to the one below (which assumes equal numbers of subjects in each group). (c) The results observed in this study would rarely have occurred by chance if vitamin C were ineffective.



3.48. The population is faculty members at Mongolian public universities, and the sample is the 300 faculty members to whom the survey was sent. Because we do not know how many responses were received, we cannot determine the response rate.

Note: We might consider the population to be either the 2500 faculty members on the list, or the larger group of “all current and future faculty members.” In the latter case, those on the list constitute the sampling frame—the subset of the population from which our sample will be selected.

The sample might be considered to be only those faculty who actually responded to the survey (rather than the 300 selected), because that is the actual group from which we “draw conclusions about the whole.”

3.49. The population is all forest owners in the region. The sample is the 772 forest owners contacted. The response rate is $\frac{348}{772} \doteq 45\%$. Aside from the given information, we would like to know the sample design (and perhaps some other things).

Note: It would also be reasonable to consider the sample to be the 348 forest owners who returned the survey, because that is the actual group from which we “draw conclusions about the whole.”

3.50. To use Table B, number the list from 0 to 9 and choose two single digits. (One can also assign labels 01–10, but that would require two-digit numbers, and we would almost certainly end up skipping over many pairs of digits before we found two in the desired range.)

It is worth noting that choosing an SRS is often described as “pulling names out of a hat.” For long lists, it is often impractical to do this literally, but with such a small list, one really could write each ringtone on a slip of paper and choose two slips at random.

3.51. See the solution to the previous exercise; for this problem, we need to choose three items instead of two, but it is otherwise the same.

3.52. (a) This is a multistage sample: We first sample three of the seven course sections, then eight from each chosen section. (b) This is an SRS: Each student has the same chance ($5/55$) of being selected. (c) This is a voluntary response sample: Only those who visit the site can participate (if they choose to). (d) This is a stratified random sample: Males and females (the strata) are sampled separately.

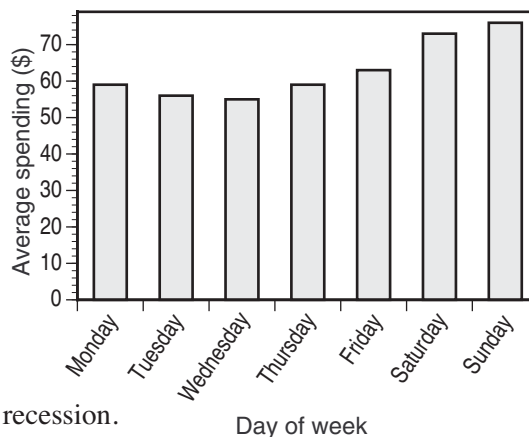
3.53. (a) This statement confuses the ideas of population and sample. (If the entire population is found in our sample, we have a *census* rather than a sample.) (b) “Dihydrogen monoxide” is H_2O . Any concern about the dangers posed by water most likely means that the respondent did not know what dihydrogen monoxide was, and was too embarrassed to admit it. (Conceivably, the respondent knew the question was about water and had concerns arising from a bad experience of flood damage or near-drowning. But misunderstanding seems to be more likely.) (c) Honest answers to such questions are difficult to obtain even in an anonymous survey; in a public setting like this, it would be surprising if there were any raised hands (even though there are likely to be at least a few cheaters in the room).

3.54. (a) The content of a single chapter is not random; choose random words from random pages. (b) Students who are registered for an early-morning class might have different characteristics from those who avoid such classes. (c) Alphabetic order is not random. One problem is that the sample might include people with the same last name (siblings, spouses, etc.). Additionally, some last names tend to be more common in some ethnic groups.

3.55. The population is (all) local businesses. The sample is the 73 businesses that return the questionnaire, or the 150 businesses selected. The nonresponse rate is $51.3\% = \frac{77}{150}$.

Note: The definition of “sample” makes it somewhat unclear whether the sample includes all the businesses selected or only those that responded. My inclination is toward the latter (the smaller group), which is consistent with the idea that the sample is “a part of the population that we actually examine.”

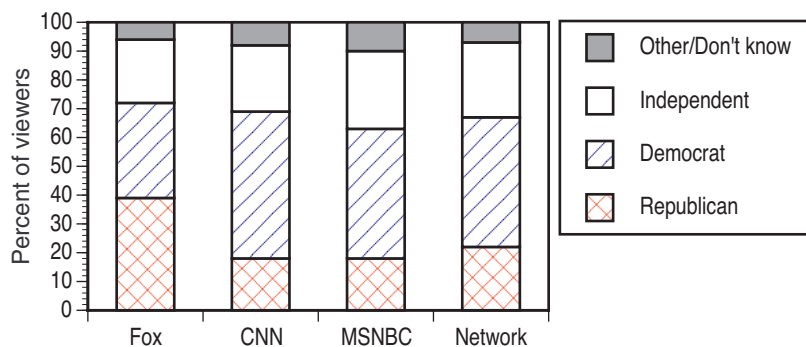
3.56. (a) Shown is a possible display (as a bar graph). A plot similar to Figure 3.7 would tell the same story—namely, that spending is highest on the weekends, and drops during the middle of the week. Ideally, bars should be in chronological order, although students might choose to start with Sunday rather than Monday. Some students might arrange bars in increasing or decreasing order of height.



(b) The exclusion of the Christmas shopping season might impact the numbers. In addition, much of this period fell during an economic recession.

3.57. Note that the numbers add to 100% down the columns; that is, 39% is the percent of Fox viewers who are Republicans, *not* the percent of Republicans who watch Fox.

Students might display the data using a stacked bar graph like the one below, or side-by-side bars. (They could also make four pie charts, but comparing slices across pies is difficult.) The most obvious observation is that the party identification of Fox's audience is noticeably different from the other three sources.



3.58. Exact descriptions of the populations may vary. **(a)** All current students—or perhaps all current students who were enrolled during the year prior to the change. (The latter would be appropriate if we want opinions based on a comparison between the new and old curricula.) **(b)** All U.S. households. **(c)** Adult residents of the United States.

3.59. Numbering from 01 to 33 alphabetically (down the columns), we enter Table B at line 137 and choose:

12 = Country View, 14 = Crestview, 11 = Country Squire, 16 = Fairington, 08 = Burberry
See note on page 50 about using Table B.

3.60. Assign labels 001 to 200. To use Table B, take three digits at a time beginning on line 112; the first five pixels are 089, 064, 032, 117, and 003.

3.61. Population = 1 to **200**, Select a sample of size **25**, then click **Reset** and **Sample**.

3.62. With the applet: Population = 1 to **371**, Select a sample of size **25**, then click **Reset** and **Sample**. With Table B, line 120 gives the codes labeled 354, 239, 193, 099, and 262.

3.63. One could use the labels already assigned to the blocks, but that would mean skipping a lot of four-digit combinations that do not correspond to any block. An alternative would be to drop the second digit and use labels 100–105, 200–211, and 300–325. But by far the simplest approach is to assign labels 01–44 (in numerical order by the four-digit numbers already assigned), enter the table at line 135, and select:

39 (block 3020), 10 (2003), 07 (2000), 11 (2004), and 20 (3001)

See note on page 50 about using Table B.

3.64. If one always begins at the same place, then the results could not really be called random.

3.65. The sample will vary with the starting line in Table B. The simplest method is to use the last digit of the numbers assigned to the blocks in Group 1 (that is, assign the labels 0–5), then choose one of those blocks; use the last two digits of the blocks in Group 2 (00–11) and choose two of those, and finally use the last two digits of the blocks in Group 3 (00–25) and choose three of them.

3.66. (a) If we choose one of the first 45 students and then every 45th name after that, we will have a total of $\frac{9000}{45} = 200$ names. **(b)** Label the first 45 names 01–45. Beginning at line 125, the first number we find is 21, so we choose names 21, 66, 111, . . .

3.67. Considering the 9000 students of Exercise 3.66, each student is equally likely; specifically, each name has chance $1/45$ of being selected. To see this, note that each of the first 45 has chance $1/45$ because one is chosen at random. But each student in the second 45 is chosen exactly when the corresponding student in the first 45 is, so each of the second 45 also has chance $1/45$. And so on.

This is not an SRS because the only possible samples have exactly one name from the first 45, one name from the second 45, and so on; that is, there are only 45 possible samples. An SRS could contain *any* 200 of the 9000 students in the population.

3.68. (a) This is a stratified random sample. **(b)** Label from 01 through 27; beginning at line 122, we choose:

13 (805), 15 (760), 05 (916), 09 (510), 08 (925),

27 (619), 07 (415), 10 (650), 25 (909), and 23 (310)

Note: The area codes are in north-south order if we read across the rows; that is how they were labeled for this solution. Students might label down rather than across; the sample should include the same set of labels but a different list of area codes.

3.69. Assign labels 01–36 for the Climax 1 group, 01–72 for the Climax 2 group, and so on.

Then beginning at line 140, choose:

12, 32, 13, 04 from the Climax 1 group and (continuing on in Table B)

51, 44, 72, 32, 18, 19, 40 from the Climax 2 group

24, 28, 23 from the Climax 3 group and

29, 12, 16, 25 from the mature secondary group

See note on page 50 about using Table B.

3.70. Label the students 01, . . . , 30 and use Table B. Then label the faculty 0, . . . , 9 and use the table again. (You could also label the faculty from 01 to 10, but that would needlessly require two-digit labels.)

Note: *Students often try some fallacious method of choosing both samples simultaneously. We simply want to choose two separate SRSs: one from the students and one from the faculty. See note on page 50 about using Table B.*

3.71. Each student has a 10% chance: 3 out of 30 over-21 students, and 2 of 20 under-21 students. This is not an SRS because not every group of 5 students can be chosen; the only possible samples are those with 3 older and 2 younger students.

3.72. Label the 500 midsize accounts from 001 to 500, and the 4400 small accounts from 0001 to 4400. On line 115, we first encounter numbers 417, 494, 322, 247, and 097 for the midsize group, then 3698, 1452, 2605, 2480, and 3716 for the small group. See note on page 50 about using Table B.

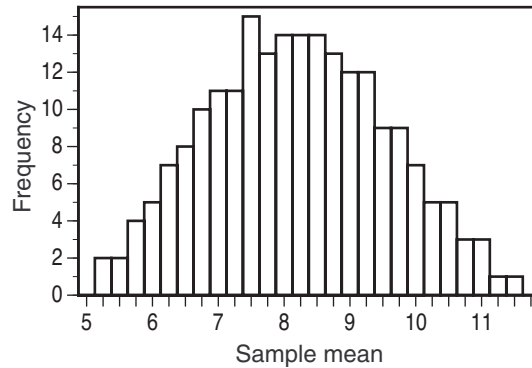
3.73. (a) This design would omit households without telephones or with unlisted numbers. Such households would likely be made up of poor individuals (who cannot afford a phone), those who choose not to have phones, and those who do not wish to have their phone numbers published. **(b)** Those with unlisted numbers would be included in the sampling frame when a random-digit dialer is used.

3.74. (a) This will almost certainly produce a positive response because it draws the dubious conclusion that cell phones *cause* brain cancer. Some people who drive cars, or eat carrots, or vote Republican develop brain cancer, too. Do we conclude that these activities should come with warning labels, also? **(b)** The phrasing of this question will tend to make people respond in favor of national health insurance: It lists two benefits of such a system, and no arguments from the other side of the issue. **(c)** This sentence is so convoluted and complicated that it is almost unreadable; it is also vague (what sort of “economic incentives”? How much would this cost?). A better phrasing might be, “Would you be willing to pay more for the products you buy if the extra cost were used to conserve resources by encouraging recycling?” That is still vague, but less so, and is written in plain English.

3.75. The first wording brought the higher numbers in favor of a tax cut; “new government programs” has considerably less appeal than the list of specific programs given in the second wording.

- 3.76.** Children from larger families will be overrepresented in such a sample. Student explanations of why will vary; a simple illustration can aid in understanding this effect. Suppose that there are 100 families with children; 60 families have one child and the other 40 have three. Then there are a total of 180 children (an average of 1.8 children per family), and *two-thirds* (120) of those children come from families with three children. Therefore, if we had a class (a sample) chosen from these 180 children, only one-third of the class would answer “one” to the teacher’s question, and the rest would say “three.” This would give an average of about 2.3 children per family.
- 3.78.** Responses to public opinion polls can be affected by things like the wording of the question, as was the case here: Both statements address the question of how to distribute wealth in a society, but subtle (and not-so-subtle) slants in the wording suggest that the public holds conflicting opinions on the subjects.
- 3.79.** The population is undergraduate college students. The sample is the 2036 students. (We assume they were randomly selected.)
- 3.80.** No; this is a voluntary response sample. The procedures described in the text apply to data gathered from an SRS.
- 3.81.** The larger sample would have less sampling variability. (That is, the results would have a higher probability of being closer to the “truth.”)
- 3.82.** (a) Parameters are associated with the population; statistics describe samples. (b) Bias means that the center of sampling distribution is not equal to the true value of the parameter; that is, bias is systematic under- or over-estimation. Variability refers to the spread (not the center) of the sampling distribution. (c) Large samples generally have lower variability, but if the samples are biased, that lower variability is of little use. (In addition, larger samples generally come at a cost; the added cost might not justify the decrease in variability.) (d) A sampling distribution might be visualized (or even simulated) with a computer, but it arises from the process of sampling, not from computation.
- 3.83.** (a) Population: college students. Sample: 17,096 students. (b) Population: restaurant workers. Sample: 100 workers. (c) Population: longleaf pine trees. Sample: 40 trees.
- 3.84.** (a) High bias, high variability (many are low, wide scatter). (b) Low bias, low variability (close to parameter, little scatter). (c) Low bias, high variability (neither too low nor too high, wide scatter). (d) High bias, low variability (too high, little scatter).
- Note:** Make sure that students understand that “high bias” means that the values are far from the parameter, not that they are too high.

- 3.85. (a)** The scores will vary depending on the starting row. Note that the smallest possible mean is 5.25 (from the sample 3, 5, 6, 7) and the largest is 11.5 (from 9, 10, 12, 15).
(b) Answers will vary. On the right is the (exact) sampling distribution, showing all possible values of the experiment (so the first rectangle is for 5.25, the next is for 5.5, etc.). Note that it looks roughly Normal; if we had taken a larger sample from a larger population, it would appear even more Normal.



Note: This histogram was found by considering all $\binom{10}{4} = 210$ of the possible samples. A collection of only 10 random samples will, of course, be considerably less detailed.

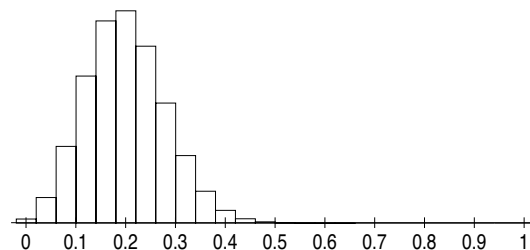
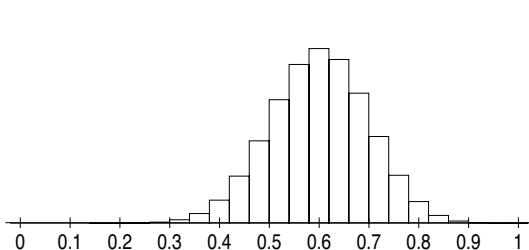
- 3.86.** No: With sufficiently large populations (“at least 100 times larger than the sample”), the variability (and margin of error) depends on the sample size.

- 3.87. (a)** This is a multistage sample. **(b)** Attitudes in smaller countries (many of which were not surveyed) might be different. **(c)** An individual country’s reported percent will typically differ from its true percent by no more than the stated margin of error. (The margins of error differ among the countries because the sample sizes were not all the same.)

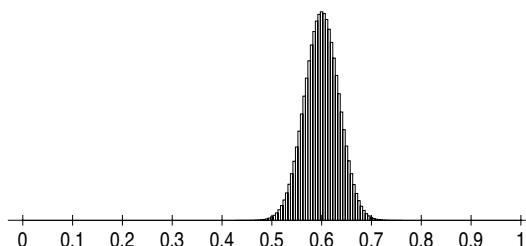
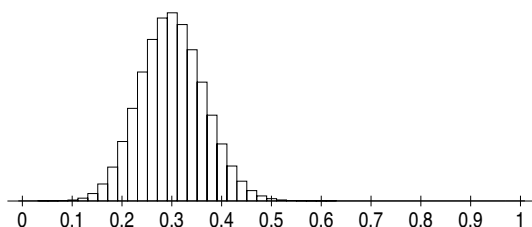
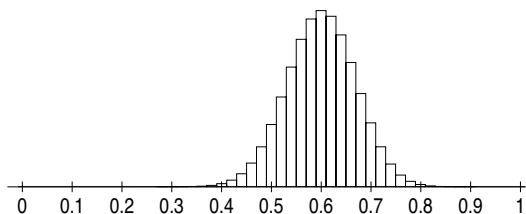
Note: The number of countries in the world is about 195 (the exact number depends on the criteria of what constitutes a separate country). That means that about 60 countries are not represented in this survey.

- 3.88. (a)** The population is Ontario residents; the sample is the 61,239 people interviewed.
(b) The sample size is very large, so if there were large numbers of both sexes in the sample—this is a safe assumption because we are told this is a “random sample”—these two numbers should be fairly accurate reflections of the values for the whole population.

- 3.89. (a)** The histogram should be centered at about 0.6 (with quite a bit of spread). For reference, the theoretical histogram is shown below on the left; student results should have a similar appearance. **(b)** The histogram should be centered at about 0.2 (with quite a bit of spread). The theoretical histogram is shown below on the right.

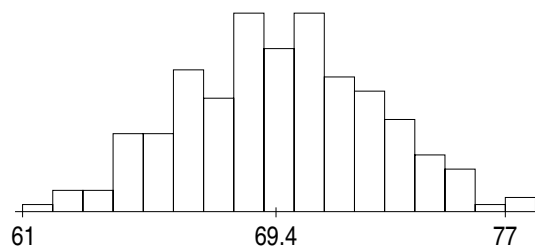
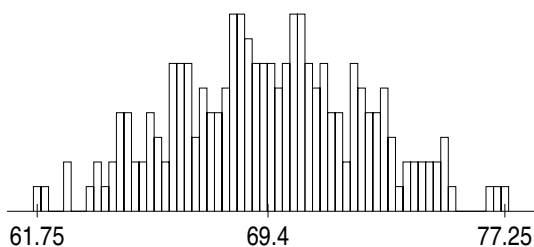


- 3.90. (a)** The histogram of this theoretical sampling distribution is shown (on the right) for reference. **(b)** This theoretical sampling distribution is shown below on the left. Students should observe that their two stemplots have clearly different centers (near 0.6 and 0.3, respectively) but similar spreads. **(c)** The theoretical sampling distribution is below on the right. Compared to the distribution of part (a), this has the same center but is about half as wide; that is, the spread is about half as much when the sample size is multiplied by 4. (The vertical scale of this graph is not the same as the other two; it should be about twice as tall as it is since it is only about half as wide.)



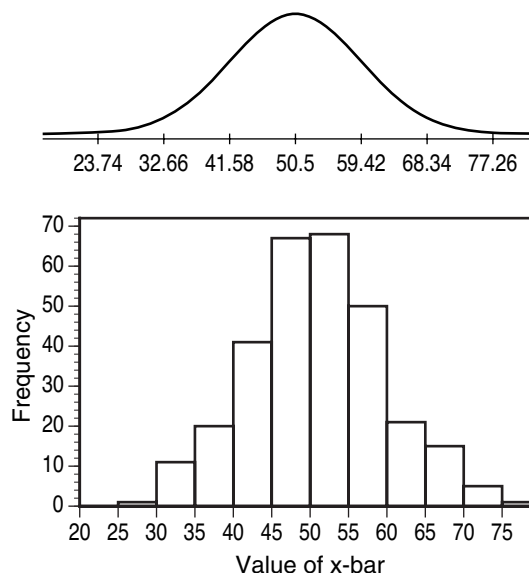
- 3.91. (a)** The scores will vary depending on the starting row. Note that the smallest possible mean is 61.75 (from the sample 58, 62, 62, 65) and the largest is 77.25 (from 73, 74, 80, 82). **(b)** Answers will vary; shown below are two views of the (exact) sampling distribution. The first shows all possible values of the experiment (so the first rectangle is for 61.75, the next is for 62.00, etc.); the other shows values grouped from 61 to 61.75, 62 to 62.75, etc. (which makes the histogram less bumpy). The tallest rectangle in the first picture is 8 units; in the second, the tallest is 28 units.

Note: These histograms were found by considering all $\binom{10}{4} = 210$ of the possible samples. It happens that half (105) of those samples yield a mean smaller than 69.4 and half yield a greater mean.



3.92. Student results will vary greatly, and ten values of \bar{x} will give little indication of the appearance of the sampling distribution. In fact, the sampling distribution of \bar{x} is approximately Normal with a mean of 50.5 and a standard deviation of about 8.92; this approximating Normal distribution is shown on the right (above). Therefore, nearly every sample of size 10 would yield a mean between 23 and 78.

The shape of the sampling distribution becomes more apparent if the results of many students are pooled. Below on the right is an example based on 300 sample means, which might arise from pooling all the results in a class of 30.



Note: Because the values in these samples are not independent (there can be no repeats), a stronger version of the central limit theorem is needed to determine that the sampling distribution is approximately Normal. Confirming the standard deviation given above is a reasonably difficult exercise even for a mathematics major.

3.93. (a) Below is the population stemplot (which gives the same information as a histogram).

The (population) mean GPA is $\mu \doteq 2.6352$, and the standard deviation is $\sigma \doteq 0.7794$.

[Technically, we should take $\sigma \doteq 0.7777$, which comes from dividing by n rather than $n - 1$, but few (if any) students would know this, and it has little effect on the results.]

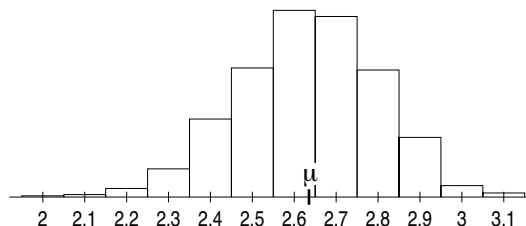
(b) & (c) Results will vary; these histograms are not shown. Not every sample of size 20 could be viewed as “generally representative of the population,” but most should bear at least some resemblance to the population distribution.

```

0 | 134
0 | 567889
1 | 0011233444
1 | 5566667888888888999999
2 | 00000000011111111222222223333333344444444
2 | 5555555555556666666677777777788888888888999999
3 | 0000000000000011111112222222233333333333344444444
3 | 55666666666677777788889
4 | 0000

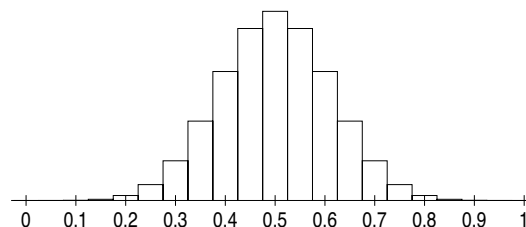
```

3.94. (a) Shown for reference is a histogram of the approximate sampling distribution of \bar{x} . This distribution is difficult to find exactly, but based on 1000 simulated samples, it is approximately Normal with mean 2.6352 (the same as μ) and standard deviation $s_{\bar{x}} \doteq 0.167$. (Therefore, \bar{x} will almost always be between 2.13 and 3.14.) **(b)** Results may vary, but most students should see no strong suggestion of bias. **(c)** Student means and standard deviations will vary, but for most (if not all) students, their values should meet the expectations (close to $\mu \doteq 2.6352$ and less than $\sigma \doteq 0.78$).



Note: Observe that the distribution of \bar{x} is slightly left-skewed, but less skewed than the population distribution. Also note that $s_{\bar{x}}$, the standard deviation of the sampling distribution, is smaller than $\sigma/\sqrt{20} \doteq 0.174$, since we are sampling without replacement.

3.95. (a) Answers will vary. If, for example, eight heads are observed, then $\hat{p} = \frac{8}{20} = 0.4 = 40\%$. **(b)** Note that all the leaves in the stemplot should be either 0 or 5 since all possible \hat{p} -values end in 0 or 5. For comparison, here is a histogram of the sampling distribution (assuming p really is 0.5). An individual student's stemplot will probably only roughly approximate this distribution, but pooled efforts should be fairly close.



Many of the questions in Section 3.4 (Ethics), Exercises 3.96–3.117, are matters of opinion and may be better used for class discussion rather than as assigned homework. A few comments are included here.

3.96. These three proposals are clearly in increasing order of risk. Most students will likely consider that (a) qualifies as minimal risk, and most will agree that (c) goes beyond minimal risk.

3.97. (a) A nonscientist might raise different viewpoints and concerns from those considered by scientists. **(b)** Answers will vary.

3.98. It is good to plainly state the purpose of the research (“To study how people’s religious beliefs and their feelings about authority are related”). Stating the research *thesis* (that orthodox religious belief are associated with authoritarian personalities) would cause bias.

3.102. (a) Ethical issues include informed consent and confidentiality; random assignment generally is not an ethical consideration. **(b)** “Once research begins, the board monitors its progress at least once a year.” **(c)** Harm need not be physical; psychological harm also needs to be considered.

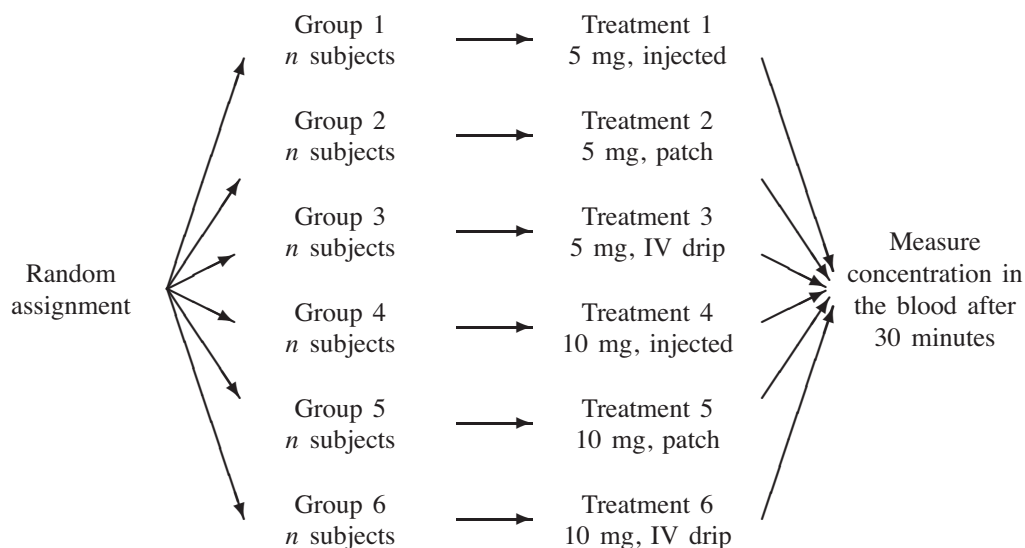
- 3.103.** To control for changes in the mass spectrometer over time, we should alternate between control and cancer samples.
- 3.105.** The articles are “Facebook and academic performance: Reconciling a media sensation with data” (Josh Pasek, eian more, Eszter Hargittai), a critique of the first article called “A response to reconciling a media sensation with data” (Aryn C. Karpinski), and a response to the critique (“Some clarifications on the Facebook-GPA study and Karpinski’s response”) by the original authors. In case these articles are not available at the address given in the text, they might be found elsewhere with a Web search.
- 3.106. (a)** The simplest approach is to label from 00001 through 14959 and then take five digits at a time from the table. A few clever students might think of some ways to make this process more efficient, such as taking the first random digit chosen as “0” if it is even and “1” if odd. (This way, fewer numbers need to be ignored.) **(b)** Using labels 00001–14959, we choose 05995, 06788, and 14293. Students who try an alternate approach may have a different sample.
- 3.108. (a)** A sample survey: We want to gather information about a population (U.S. residents) based on a sample. **(b)** An experiment: We want to establish a cause-and-effect relationship between teaching method and amount learned. **(c)** An observational study: There is no particular population from which we will sample; we simply observe “your teachers,” much like an animal behavioral specialist might study animals in the wild.
- 3.111.** They cannot be anonymous because the interviews are conducted in person in the subject’s home. They are certainly kept confidential.
Note: For more information about this survey, see the GSS Web site:
www.norc.org/GSS+Website
- 3.112.** This offers anonymity, since names are never revealed. (However, faces are seen, so there may be some chance of someone’s identity becoming known.)
- 3.116. (a)** Those being surveyed should be told the kind of questions they will be asked and the approximate amount of time required. **(b)** Giving the name and address of the organization may give the respondents a sense that they have an avenue to complain should they feel offended or mistreated by the pollster. **(c)** At the time that the questions are being asked, knowing who is paying for a poll may introduce bias, perhaps due to nonresponse (not wanting to give what might be considered a “wrong” answer). When information about a poll is made public, though, the poll’s sponsor should be announced.
- 3.120.** At norc.org, search for “Consumer Finances” or “SCF,” and from the SCF page, click on the link to “website for SCF respondents.” At the time this manual was written, the pledge was found at www.norc.org/scf2010/Confidentiality.html.
- 3.121. (a)** You need information about a random selection of his games, not just the ones he chooses to talk about. **(b)** These students may have chosen to sit in the front; all students should be randomly assigned to their seats.

- 3.122.** (a) A matched pairs design (two halves of the same board would have similar properties). (b) A sample survey (with a stratified sample: smokers and nonsmokers). (c) A block design (blocked by gender).
- 3.123.** This is an experiment because each subject is (randomly, we assume) assigned to a treatment. The explanatory variable is the price history seen by the subject (steady prices or fluctuating prices), and the response variable is the price the subject expects to pay.
- 3.124.** (a) For example, one could select (or recruit) a sample and assess each person's calcium intake (perhaps by having them record what they eat for a week), and measure his/her TBBMD. (b) For example, measure each subject's TBBMD, then randomly assign half the subjects to take a calcium supplement, and the other half to take a placebo. After a suitable period, measure TBBMD again. (c) The experiment, while more complicated, gives better information about the relationship between these variables, because it controls for other factors that may affect bone health.
- 3.126.** Each subject should taste both kinds of fries in a randomly selected order and then be asked about preference. One question to consider is whether they should have ketchup available; many people typically eat fries with ketchup, and its presence or absence might affect their preferences. If ketchup is used, should one use the same ketchup for both, or a sample of the ketchup from each restaurant?
- 3.127.** The two factors are gear (three levels) and steepness of the course (number of levels not specified). Assuming there are at least three steepness levels—which seems like the smallest reasonable choice—that means at least nine treatments. Randomization should be used to determine the order in which the treatments are applied. Note that we must allow ample recovery time between trials, and it would be best to have the rider try each treatment several times.
- 3.129.** (a) One possible population: all full-time undergraduate students in the fall term on a list provided by the registrar. (b) A stratified sample with 125 students from each year is one possibility. (c) Mailed (or emailed) questionnaires might have high nonresponse rates. Telephone interviews exclude those without phones and may mean repeated calling for those who are not home. Face-to-face interviews might be more costly than your funding will allow. There might also be some response bias: Some students might be hesitant about criticizing the faculty (while others might be far too eager to do so).

3.130. (a) For the two factors (administration method, with three levels, and dosage, with two levels), the treatment combinations are shown in the table on the right, and the design is diagrammed below.

	Injection	Skin patch	IV drip
5 mg	1	2	3
10 mg	4	5	6

(b) Larger samples give more information; in particular, with large samples, we reduce the variability in the observed mean concentrations so that we can have more confidence that the differences we might observe are due to the treatment applied rather than random fluctuation.



3.131. Use a block design: Separate men and women, and randomly allocate each gender among the six treatments.

The remaining exercises relate to the material of Section 3.4 (Ethics). Answers are given for the first two; the rest call for student opinions, or information specific to the student's institution.

3.132. Parents who fail to return the consent form may be more likely to place less priority on education and therefore may give their children less help with homework, and so forth. Including those children in the control group is likely to lower that group's score.

Note: *This is a generalization, to be sure: We are not saying that every such parent does not value education, only that the percentage of this group that highly values education will almost certainly be lower than that percentage of the parents who return the form.*

3.133. The latter method (CASI) will show a higher percentage of drug use because respondents will generally be more comfortable (and more assured of anonymity) about revealing embarrassing or illegal behavior to a computer than to a person, so they will be more likely to be honest.