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 computer programmer (and his friends) are not representative of all young people.

3.4. Anecdotal evidence is rarely reliable for drawing conclusions. One can hardly guarantee that there would always be a soft place to land in the event of a crash; Herman's friend would likely not have fared as well had he or she landed on the pavement.

3.7. This is an observational study: No treatment was assigned to the subjects; we merely observed cell phone usage (and presence/absence of cancer). The explanatory variable is cell phone usage, and the response variable is whether or not a subject has brain cancer.

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.

3.11. 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.

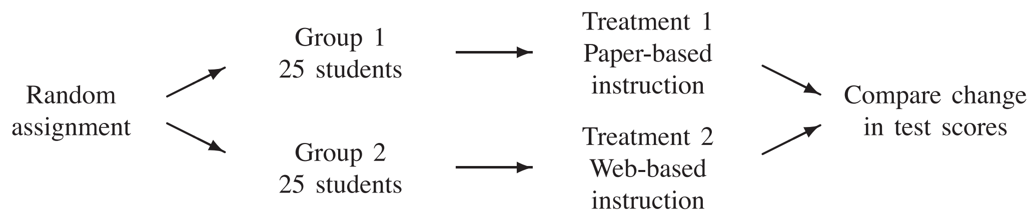
It is hard to say how much this will generalize; it seems likely that different lipids react to radiation differently.

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.

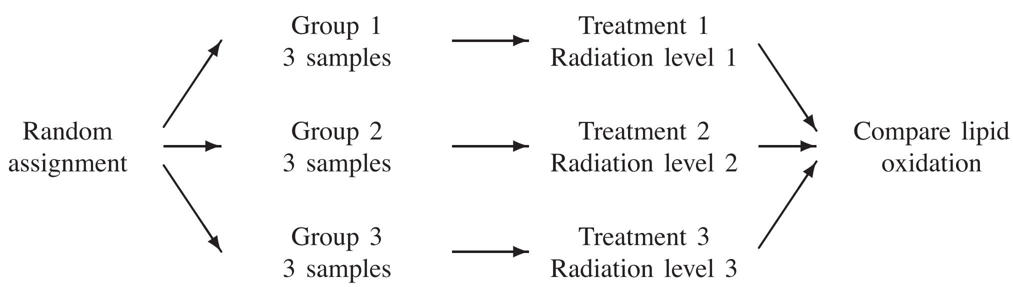
The results of this experiment should generalize to other classes (on the same topic) taught by the same instructor, but might not apply to other subject matter, or to classes taught by other instructors.

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.



3.17. **(a)** Students in the front rows have a different classroom experience from those in the back. (And if they chose their own seats, those who choose seats in the front may be different from those who choose back seats.) **(b)** There is no control group, so we have nothing to which we can compare the observed score change. **(c)** It is hard to compare different classes (zoology and botany) in different semesters.

3.18. **(a)** One problem is that people with the same last name will likely be assigned to the same group. Additionally, some last names tend to be more common in some ethnic groups; it is possible, for example, that a large proportion of the last 10 subjects might be Asian (with names like Zheng, Yang, etc.). **(b)** This randomization will not necessarily divide the subjects into two groups of four. (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 20 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 five rats from each batch to each treatment.)

3.19. Those evaluating the exams should not know which teaching approach was used, and the students should not be told that they are being taught using the new (or old) method.

3.20. For example, we might block by gender, by year in school, or by housing type (dorm/off-campus/Greek).

3.21. There should be at least two treatments (water and compost tea), and possibly more (such as different volumes or concentrations of compost tea). A complete description should include the number of units (plants) assigned to each treatment. Possible response variables include increase in weight or height, number of leaves, etc.

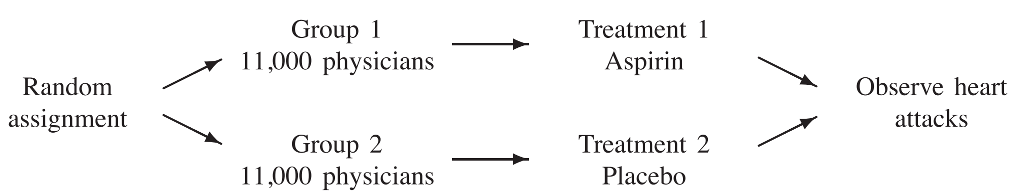
3.22. Students might envision different treatments; one possibility is to have some volunteers go through a training session, while others are given a written set of instructions, or watch a video. For the response variable(s), we need some measure of training effectiveness; perhaps we could have the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

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

3.24. 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.

3.25. 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.

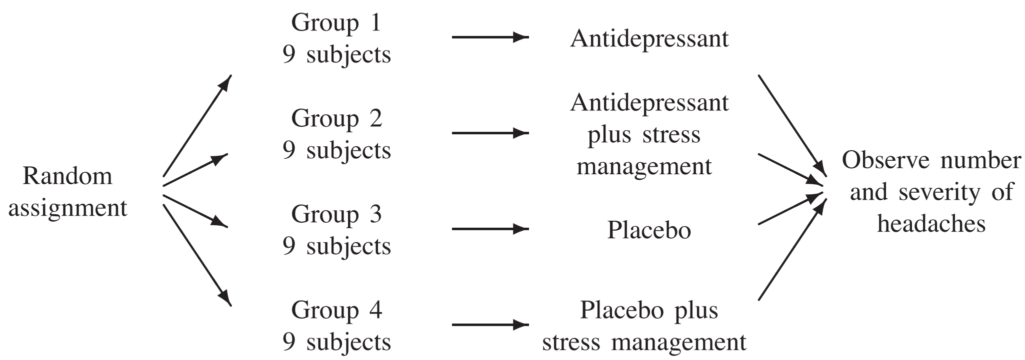
3.26. **(a)** The subjects are the physicians, the factor is medication (aspirin or placebo), and the response variable is health, specifically whether the subjects have heart attacks. **(b)** Below. **(c)** The difference in the number of heart attacks between the two groups was so great that it would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



3.27. 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	

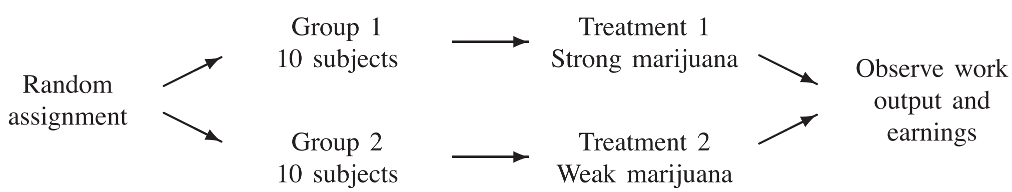
The other nine subjects are in Group 4. See note on page 51 about using Table B. (In particular, be aware that in the Student's Study Guide, the solution to this problem is found using labels 00–35.)



3.28. **(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

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 51 about using Table B.

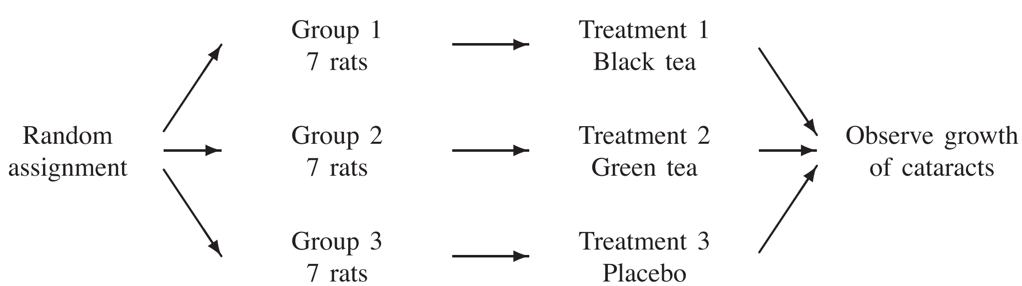


3.29. **(a)** Diagram below. **(b)** If we assign labels 01, . . . , 21 and begin on line 120, then we select:

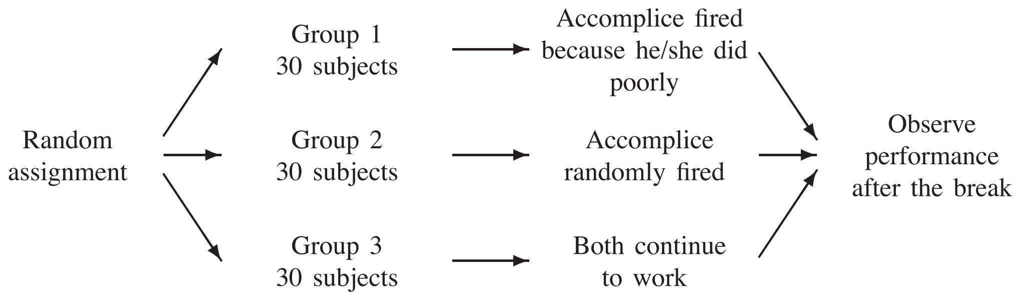
16, 04, 21, 19, 07, 10, and 13 for Group 1

15, 05, 09, 08, 18, 03, and 01 for Group 2

The remaining rats are assigned to the placebo group. See note on page 51 about using Table B.



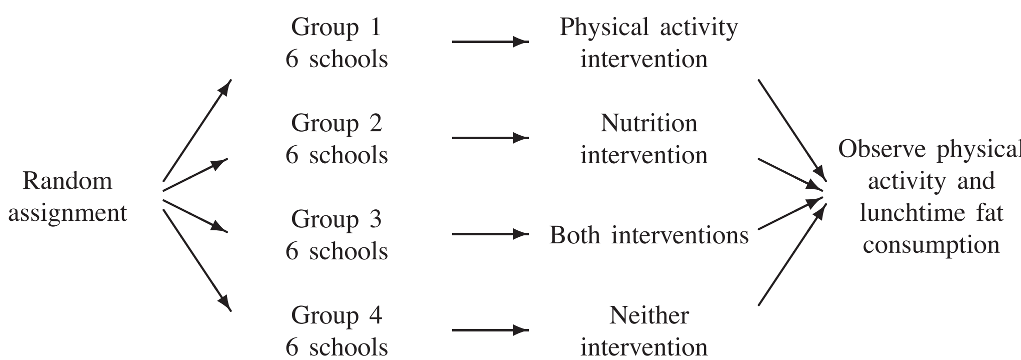
3.30. (a) Diagram below. **(b)** Using line 153 from Table B, the first four subjects are 07, 88, 65, and 68. See note on page 51 about using Table B.



3.31. 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

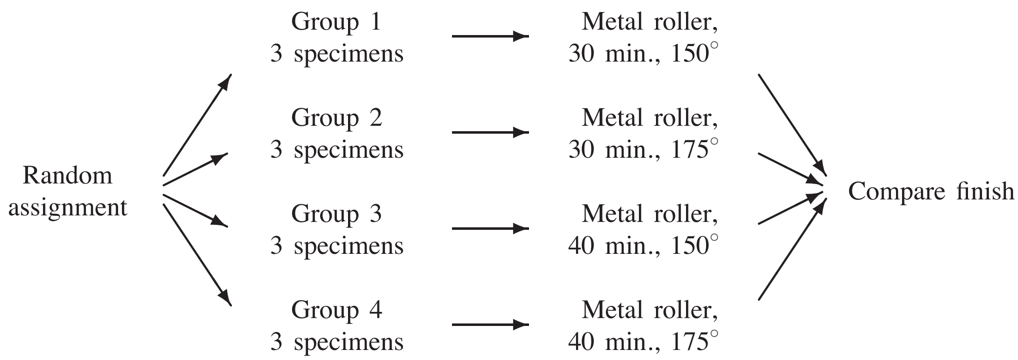
The rest are assigned to Group 4. See note on page 51 about using Table B.



3.32. (a) The table below shows the 16 treatments—four levels for each of the two factors. **(b)** A diagram is not shown here (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

3.33. **(a)** There are three factors (roller type, dyeing cycle time, and temperature), each with two levels, for a total of $2^3 = 8$ treatments. The experiment therefore requires 24 fabric specimens. **(b)** In the interest of space, only the top half of the diagram is shown below. The other half consists of Groups 5 to 8, for which the treatments have natural bristle rollers instead of metal rollers.



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

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

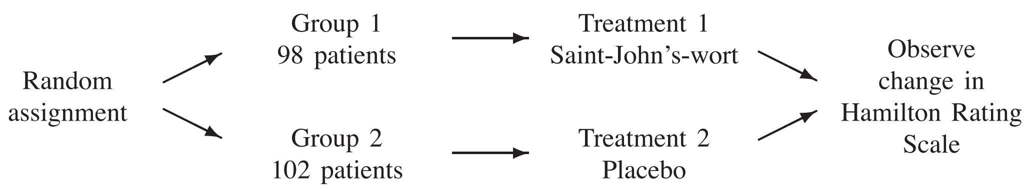
(b) Without resetting, click Sample again.

3.36. 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.

3.37. The first design 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). The second design 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.

3.38. “Randomized” means that patients were randomly assigned to receive either the standard morphine treatment or CR morphine tablets. “Double blind” means that the treatment assigned to a patient was unknown to both the patient and those responsible for assessing the effectiveness of that treatment. (It is not clear how the treatment was hidden from the patients because they would know when they received the morphine.) “Comparative” means researchers compared the effectiveness of two treatments, rather than simply trying to assess the effectiveness of one treatment—that is, researchers did not simply change over to CR morphine and try to judge if it was better than the standard treatment had been in the past.

3.39. **(a)** "Randomized" means that patients were randomly assigned to receive either Saint-John's-wort or a placebo. "Double blind" means that the treatment assigned to a patient was unknown to both the patient and those responsible for assessing the effectiveness of that treatment. "Placebo-controlled" means that some of the subjects were given placebos. Even though these possess no medical properties, some subjects may show improvement or benefits just as a result of participating in the experiment; the placebos allow those doing the study to observe this effect. **(b)** Diagram below.



3.40. The mean Monday return for the first three weeks of the month was both different from zero and higher than the mean for the last two Mondays. However, the difference from zero was small enough that it might have occurred purely by chance (and so it gives no reason to suspect that the first three Mondays tend to produce negative returns). On the other hand, the difference between the first three Mondays and the last two Mondays was so large that it would rarely occur by chance, leading us to conclude that the last two Mondays really do (for whatever reason) tend to yield lower returns than the first three Mondays.

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* $P(7 \le X \le 13) \doteq 0.974$. *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.

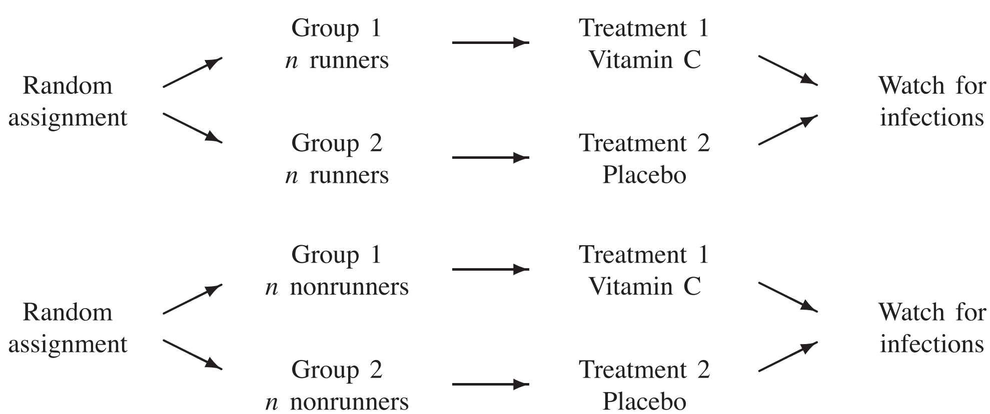
If we label the players 01 through 20 and begin on line 140, we choose 12, 13, 04, 18, 19, 16, 02, 08, 17, 10 to be in the oxygen-first group. See note on page 51 about using Table B. (In particular, be aware that solution given in the Student's Study Guide is different from that described here.)

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 those who responded to the survey (up to 300 faculty members). Because we do not know how many responses were received, we cannot determine the response rate.

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

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).

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 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).

3.53. (a) The content of a single chapter is not random; choose random words from random pages. **(b)** Students who are registered for a 7:30 class might have different characteristics from those who avoid such classes. **(c)** Alphabetic order is not random. For example, some last names occur more often in some ethnic groups—see the solution to Exercise 3.18(a).

3.54. 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."*

3.55. **(a)** Population: U.S. adults. Sample size: 1001. **(b)** Note that polls like this sometimes report only results for "those expressing an opinion." One can argue for either approach.

3.56. 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.

3.57. 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 51 about using Table B.

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

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

3.60. 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.

3.61. 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 51 about using Table B.

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

3.63. 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.

3.64. **(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, ...

3.65. Considering the 9000 students of Exercise 3.64, 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.

3.66. **(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.*

3.67. 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 51 about using Table B.

3.68. 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 51 about using Table B.*

3.69. 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.

3.70. 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 51 about using Table B.

3.71. The higher no-answer was probably the second period—more families are likely to be gone for vacations, and so on. Nonresponse of this type might underrepresent those who are more affluent (and are able to travel). In general, high nonresponse rates always make results less reliable because we do not know what information we are missing.

3.72. **(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.

3.73. **(a)** There were 1260 responses. (Note that we have no guarantee that these came from 1260 distinct people; some may have voted more than once.) **(b)** The percents in each group are $\frac{631}{1260} \doteq 50.08\%$ "yes," $\frac{564}{1260} \doteq 44.76\%$ "no," and $\frac{65}{1260} \doteq 5.16\%$ "not sure." **(c)** This voluntary response sample collects only the opinions of those who visit this site and feel strongly enough to respond.

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* 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.

3.82. **(a)** The sampling distribution describes the variation of the characteristic of a *sample*. A characteristic of a *population* does not vary; it is a fixed number. **(b)** Bias and variability are independent; any combination of high/low bias and high/low variability is possible (as Figure 3.14 illustrates). **(c)** For a given population, variability decreases with increasing sample size, so the variability for the large sample will be smaller.

3.83. **(a)** Population: college students. Sample: 17,096 students. **(b)** Population: restaurant workers. Sample: 100 workers. **(c)** Population: longleaf pine trees. Sample: 584 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*.

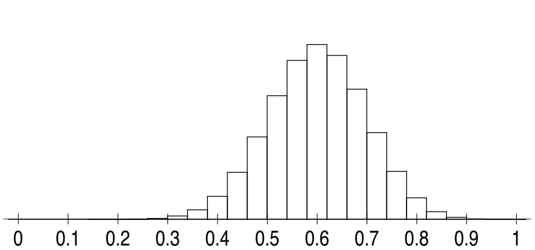
3.85. **(a)** The sample size for Hispanics was smaller. Smaller sample sizes give less information about the population and therefore lead to larger margins of error (with the same confidence level). **(b)** The sample size was so small, and the margin of error so large, that the results could not be viewed as an accurate reflection of the population of Cubans.

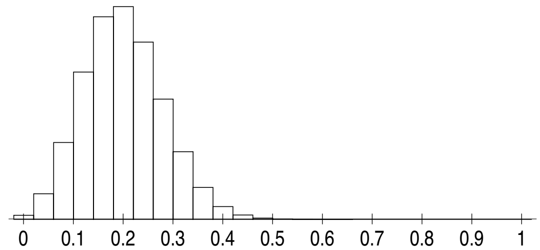
3.86. No: With sufficiently large populations ("at least 100 times larger than the sample"), the variability (and margin of error) depends on the sample size.

3.87. **(a)** Because the smallest population is still more than 100 times the sample size, the variability will be (approximately) the same for all states. **(b)** Yes, it will change—the sample size would vary from 500 in Wyoming to 35,000 in California, so the margin of error would be smaller in larger states.

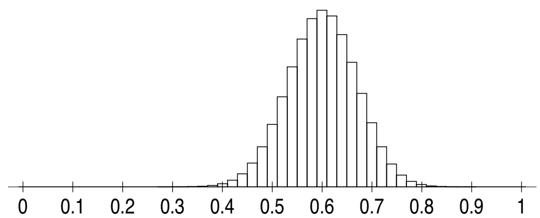
3.88. **(a)** The population is Ontario residents; the sample is the 61,239 people interviewed. **(b)** The sample size is very large, so if there were large numbers of both sexes in the sample—this is a safe assumption because we are told this is a "random sample"—these two numbers should be fairly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

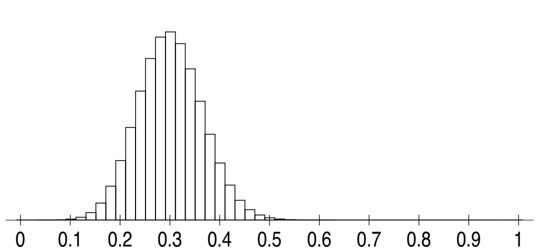
3.89. (a) The histogram should be centered at about 0.6 (with quite a bit of spread). For reference, the theoretical histogram is shown below on the left; student results should have a similar appearance. **(b)** The histogram should be centered at about 0.2 (with quite a bit of spread). The theoretical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

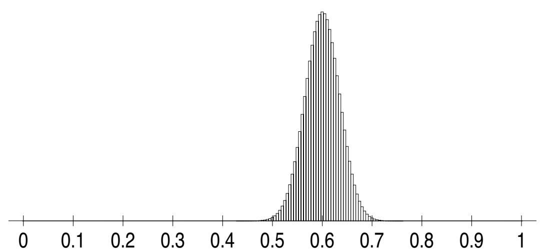




3.90. (a) The histogram of this theoretical sampling distribution is shown (on the right) for reference. **(b)** This theoretical sampling distribution is shown below on the left. Students should observe that their two stemplots have clearly different centers (near 0.6 and 0.3, respectively) but similar spreads. **(c)** The theoretical sampling distribution is below on the right. Compared to the distribution of (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.)

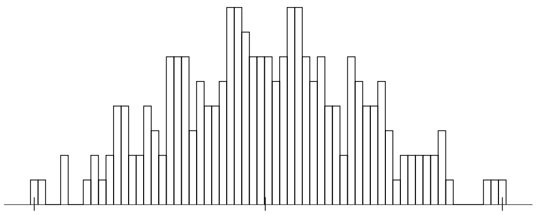


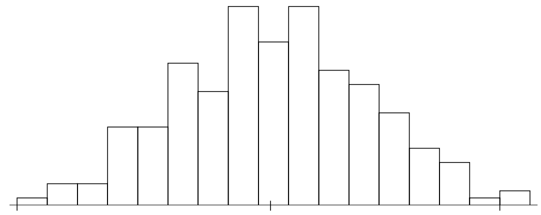




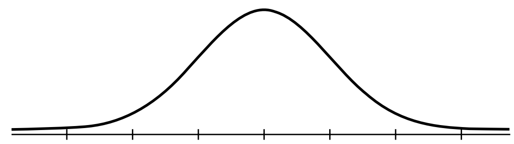
3.91. (a) The scores will vary depending on the starting row. Note that the smallest possible mean is 61.75 (from the sample 58, 62, 62, 65) and the largest is 77.25 (from 73, 74, 80, 82). **(b)** Answers will vary; shown below are two views of the (exact) sampling distribution. The first shows all possible 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.

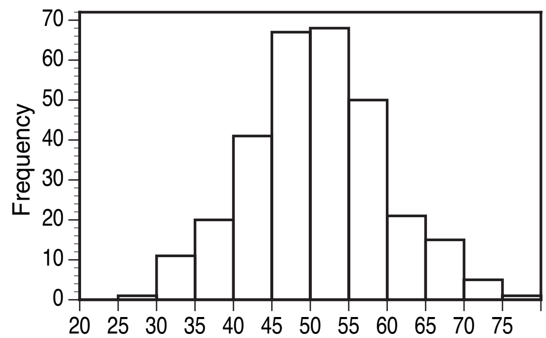
Note: *These histograms were found by considering all* $\binom{10}{4} = 210$ *of the possible* *samples. It happens that half (105) of those samples yield a mean smaller than 69.4 and half yield a greater mean.*





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.





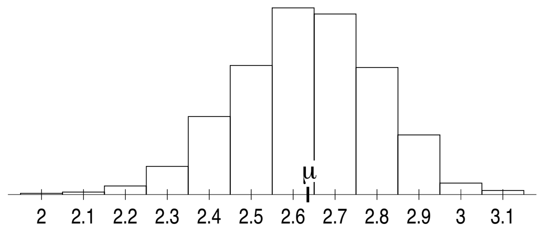
The shape of the sampling distribution becomes more apparent if the results of many students are pooled. Below on the right is an example based on 300 sample means, which might arise from pooling all the results in a class of 30.

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.

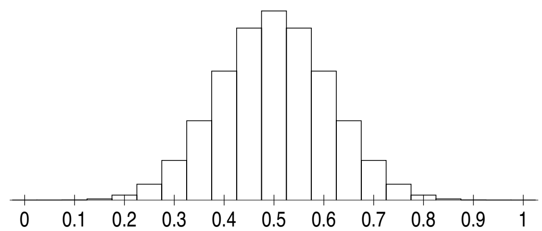
0	134
0	567889
1	0011233444
1	5566666788888888899999
2	0000000001111111112222222222333333333444444444
2	55555555555555666666667777777777777788888888888888999999
3	00000000000001111111111222222222233333333333333334444444444
3	556666666667777788889
4	0000

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 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 3.96–3.112, 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.

3.105. 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/projects/General+Social+Survey.htm`

3.106. This offers anonymity, since names are never revealed. (However, faces are seen, so there may be some chance of someone's identity becoming known.)

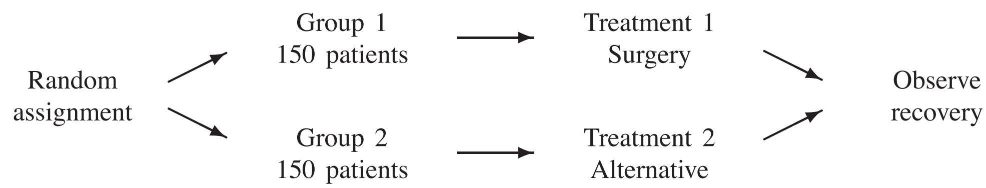
3.110. **(a)** Those being surveyed should be told the kind of questions they will be asked and the approximate amount of time required. **(b)** Giving the name and address of the organization may give the respondents a sense that they have an avenue to complain should they feel offended or mistreated 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.

3.113. **(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 03638, 07871, and 12193. Students who try an alternate approach may have a different sample.

3.114. **(a)** Possible response variables: Whether or not a subject has a job within some period of time, number of hours worked during some period, length of time before subject became employed. For the design, randomly assign about one-third of the group (3,355 subjects) to each treatment, and observe the chosen response variables after a suitable amount of time. **(b)** The simplest approach is to label from 00001 through 10065 and then take five digits at a time from the table. (This means we have to skip about 90% of the five-digit sets, as we can use only those beginning with 0, and a few beginning with 1.) With this approach, we choose 00850, 02182, and 00681 (the last of these is on line 172). More efficient labelings are possible and will lead to different samples.

3.115. **(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).

3.116. **(a)** In a serious case, when the patient has little chance of surviving, a doctor might choose not to recommend surgery; it might be seen as an unnecessary measure, bringing expense and a hospital stay with little benefit to the patient. **(b)** Diagram below.



3.117. 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.

3.118. **(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.

3.120. Each subject should taste both kinds of cheeseburger in a randomly selected order and then be asked about preference. Both burgers should have the same "fixings" (ketchup, mustard, etc.). Because some subjects might be able to identify the cheeseburgers by appearance, one might need to take additional steps (such as blindfolding or serving only the center part of the burger) in order to make this a true "blind" experiment.

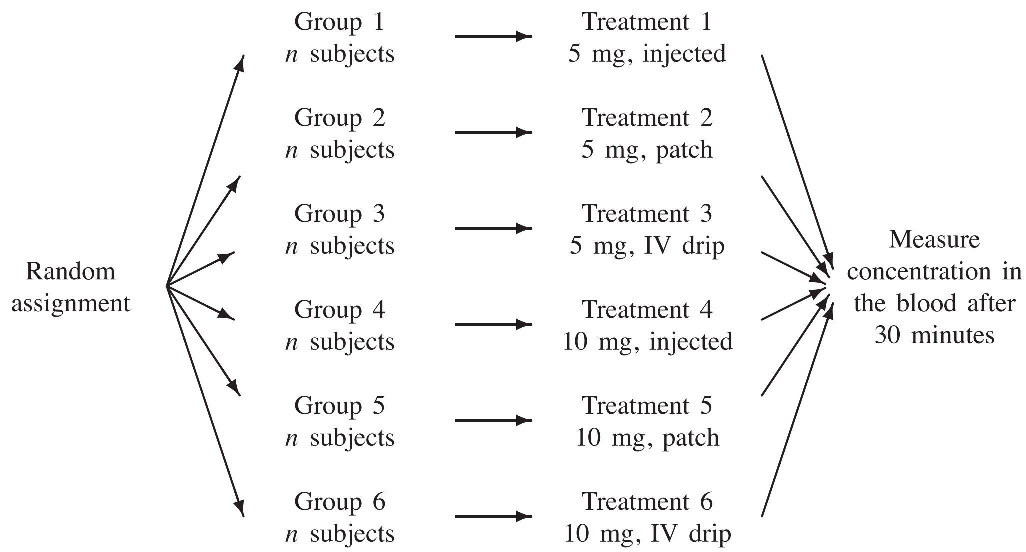
3.121. The two factors are gear (three levels) and steepness of the course (number of levels not specified). Assuming there are at least three steepness levels—which seems like the smallest reasonable choice—that means at least nine treatments. Randomization should be used to determine the order in which 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.

3.123. **(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).

3.124. **(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.



3.125. **(a)** The factors are storage method (three levels: fresh, room temperature for one month, refrigerated for one month) and preparation method (two levels: cooked immediately, or after one hour). There are therefore six treatments (summarized in the table on the right). The response variables are the tasters' color and flavor ratings. **(b)** Randomly allocate *n* potatoes to each of the six groups, then compare ratings. (Diagram not shown.) **(c)** For each taster, randomly choose the order in which the fries are tasted.

	Cooked immediately	Wait one hour
Fresh	1	2
Stored	3	4
Refrigerated	5	6

3.126. 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.

3.127. 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.*

3.128. 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.