EXERCISES

1. **Roper.** Through their *Roper Reports Worldwide*, GfK Roper conducts a global consumer survey to help multinational companies understand different consumer attitudes throughout the world. Within 30 countries, the researchers interview 1000 people aged 13–65. Their samples are designed so that they get 500 males and 500 females in each country. (www.gfkamerica.com)
   a) Are they using a simple random sample? Explain.
   b) What kind of design do you think they are using?

2. **Student Center Survey.** For their class project, a group of Statistics students decide to survey the student body to assess opinions about the proposed new student center. Their sample of 200 contained 50 first-year students, 50 sophomores, 50 juniors, and 50 seniors.
   a) Do you think the group was using an SRS? Why?
   b) What sampling design do you think they used?

3. **Emoticons.** The Web site www.gamefaqs.com asked, as their question of the day to which visitors to the site were invited to respond, *"Do you ever use emoticons when you type online?"* Of the 87,262 respondents, 27% said that they did not use emoticons. ;-(
   a) What kind of sample was this?
   b) How much confidence would you place in using 27% as an estimate of the fraction of people who use emoticons?

4. **Drug tests.** Major League Baseball tests players to see whether they are using performance-enhancing drugs. Officials select a team at random, and a drug-testing crew shows up unannounced to test all 40 players on the team. Each testing day can be considered a study of drug use in Major League Baseball.
   a) What kind of sample is this?
   b) Is that choice appropriate?

5. **Gallup.** At its Web site (www.gallup.com) the Gallup Poll publishes results of a new survey each day. Scroll down to the end, and you'll find a statement that includes words such as these:

   Results are based on telephone interviews with 1,008 national adults, aged 18 and older, conducted April 2–5, 2007. . . . In addition to sampling error, question wording and practical difficulties in conducting surveys can introduce error or bias into the findings of public opinion polls.

   a) For this survey, identify the population of interest.
   b) Gallup performs its surveys by phoning numbers generated at random by a computer program. What is the sampling frame?
   c) What problems, if any, would you be concerned about in matching the sampling frame with the population?

6. **Gallup World.** At its Web site (www.gallupworldpoll.com) the Gallup World Poll describes their methods. After one report they explained:

   Results are based on face-to-face interviews with randomly selected national samples of approximately 1,000 adults, aged 15 and older, who live permanently in each of the 21 sub-Saharan African nations surveyed. Those countries include Angola (areas where land mines might be expected were excluded), Benin, Botswana, Burkina Faso, Cameroon, Ethiopia, Ghana, Kenya, Madagascar (areas where interviewers had to walk more than 20 kilometers from a road were excluded), Mali, Mozambique, Niger, Nigeria, Senegal, Sierra Leone, South Africa, Tanzania, Togo, Uganda (the area of activity of the Lord's Resistance Army was excluded from the survey), Zambia, and Zimbabwe. . . . In all countries except Angola, Madagascar, and Uganda, the sample is representative of the entire population.

   a) Gallup is interested in sub-Saharan Africa. What kind of survey design are they using?
   b) Some of the countries surveyed have large populations. (Nigeria is estimated to have about 130 million people.) Some are quite small. (Togo's population is estimated at 5.4 million.) Nonetheless, Gallup sampled 1000 adults in each country. How does this affect the precision of its estimates for these countries?

7–14. What did they do? *For the following reports about statistical studies, identify the following items (if possible). If you can't tell, then say so—this often happens when we read about a survey.*

   a) The population
   b) The population parameter of interest
   c) The sampling frame
   d) The sample
   e) The sampling method, including whether or not randomization was employed
   f) Any potential sources of bias you can detect and any problems you see in generalizing to the population of interest

7. Consumers Union asked all subscribers whether they had used alternative medical treatments and, if so, whether they had benefited from them. For almost all of the treatments, approximately 20% of those responding reported cures or substantial improvement in their condition.

8. A question posted on the Lycos Web site on 18 June 2000 asked visitors to the site to say whether they thought that marijuana should be legally available for medicinal purposes. (www.lycos.com)

9. Researchers waited outside a bar they had randomly selected from a list of such establishments. They stopped every 10th person who came out of the bar and asked whether he or she thought drinking and driving was a serious problem.

10. Hoping to learn what issues may resonate with voters in the coming election, the campaign director for a mayoral candidate selects one block from each of the city's election districts. Staff members go there and interview all the residents they can find.

11. The Environmental Protection Agency took soil samples at 16 locations near a former industrial waste dump and checked each for evidence of toxic chemicals. They found no elevated levels of any harmful substances.

12. State police set up a roadblock to estimate the percentage of cars with up-to-date registration, insurance, and safety inspection stickers. They usually find problems with about 10% of the cars they stop.

13. A company packaging snack foods maintains quality control by randomly selecting 10 cases from each day's production and weighing the bags. Then they open one bag from each case and inspect the contents.

14. Dairy inspectors visit farms unannounced and take samples of the milk to test for contamination. If the milk is found to contain dirt, antibiotics, or other foreign matter, the milk will be destroyed and the farm reinspected until purity is restored.

15. Mistaken poll. A local TV station conducted a "PulsePoll" about the upcoming mayoral election. Evening news viewers were invited to phone in their votes, with the results to be announced on the late-night news. Based on the phone calls, the station predicted that Amabo would win the election with 52% of the vote. They were wrong: Amabo lost, getting only 46% of the vote. Do you think the station's faulty prediction is more likely to be a result of bias or sampling error? Explain.

16. Another mistaken poll. Prior to the mayoral election discussed in Exercise 15, the newspaper also conducted a poll. The paper surveyed a random sample of registered voters stratified by political party, age, sex, and area of residence. This poll predicted that Amabo would win the election with 52% of the vote. The newspaper was wrong: Amabo lost, getting only 46% of the vote. Do you think the newspaper's faulty prediction is more likely to be a result of bias or sampling error? Explain.

17. Parent opinion, part 1. In a large city school system with 20 elementary schools, the school board is considering the adoption of a new policy that would require elementary students to pass a test in order to be promoted to the next grade. The PTA wants to find out whether parents agree with this plan. Listed below are some of the ideas proposed for gathering data. For each, indicate what kind of sampling strategy is involved and what (if any) biases might result.
   a) Put a big ad in the newspaper asking people to log their opinions on the PTA Web site.
   b) Randomly select one of the elementary schools and contact every parent by phone.
   c) Send a survey home with every student, and ask parents to fill it out and return it the next day.
   d) Randomly select 20 parents from each elementary school. Send them a survey, and follow up with a phone call if they do not return the survey within a week.

18. Parent opinion, part 2. Let's revisit the school system described in Exercise 17. Four new sampling strategies have been proposed to help the PTA determine whether parents favor requiring elementary students to pass a test in order to be promoted to the next grade. For each, indi-

cate what kind of sampling strategy is involved and what (if any) biases might result.
   a) Run a poll on the local TV news, asking people to dial one of two phone numbers to indicate whether they favor or oppose the plan.
   b) Hold a PTA meeting at each of the 20 elementary schools, and tally the opinions expressed by those who attend the meetings.
   c) Randomly select one class at each elementary school and contact each of those parents.
   d) Go through the district's enrollment records, selecting every 40th parent. PTA volunteers will go to those homes to interview the people chosen.

19. Churches. For your political science class, you'd like to take a survey from a sample of all the Catholic Church members in your city. A list of churches shows 17 Catholic churches within the city limits. Rather than try to obtain a list of all members of all these churches, you decide to pick 3 churches at random. For those churches, you'll ask to get a list of all current members and contact 100 members at random.
   a) What kind of design have you used?
   b) What could go wrong with your design?

20. Playground. Some people have been complaining that the children's playground at a municipal park is too small and is in need of repair. Managers of the park decide to survey city residents to see if they believe the playground should be rebuilt. They hand out questionnaires to parents who bring children to the park. Describe possible biases in this sample.

21. Roller coasters. An amusement park has opened a new roller coaster. It is so popular that people are waiting for up to 3 hours for a 2-minute ride. Concerned about how patrons (who paid a large amount to enter the park and ride on the rides) feel about this, they survey every 10th person on the line for the roller coaster, starting from a randomly selected individual.
   a) What kind of sample is this?
   b) What is the sampling frame?
   c) Is it likely to be representative?

22. Playground, act two. The survey described in Exercise 20 asked,

Many people believe this playground is too small and in need of repair. Do you think the playground should be repaired and expanded even if that means raising the entrance fee to the park?

Describe two ways this question may lead to response bias.

23. Wording the survey. Two members of the PTA committee in Exercises 17 and 18 have proposed different questions to ask in seeking parents' opinions.

Question 1: Should elementary school–age children have to pass high-stakes tests in order to remain with their classmates?
Question 2: Should schools and students be held accountable for meeting yearly learning goals by testing students before they advance to the next grade?

   a) Do you think responses to these two questions might differ? How? What kind of bias is this?
   b) Propose a question with more neutral wording that might better assess parental opinion.

24. Banning ephedra. An online poll at a Web site asked:

A nationwide ban of the diet supplement ephedra went into effect recently. The herbal stimulant has been linked to 155 deaths and many more heart attacks and strokes. Ephedra manufacturer NVE Pharmaceuticals, claiming that the FDA lacked proof that ephedra is dangerous if used as directed, was denied a temporary restraining order on the ban yesterday by a federal judge. Do you think that ephedra should continue to be banned nationwide?

65% of 17,303 respondents said "yes." Comment on each of the following statements about this poll:

a) With a sample size that large, we can be pretty certain we know the true proportion of Americans who think ephedra should be banned.

b) The wording of the question is clearly very biased.

c) The sampling frame is all Internet users.

d) Results of this voluntary response survey can't be reliably generalized to any population of interest.

25. Survey questions. Examine each of the following questions for possible bias. If you think the question is biased, indicate how and propose a better question.

a) Should companies that pollute the environment be compelled to pay the costs of cleanup?

b) Given that 18-year-olds are old enough to vote and to serve in the military, is it fair to set the drinking age at 21?

26. More survey questions. Examine each of the following questions for possible bias. If you think the question is biased, indicate how and propose a better question.

a) Do you think high school students should be required to wear uniforms?

b) Given humanity's great tradition of exploration, do you favor continued funding for space flights?

27. Phone surveys. Anytime we conduct a survey, we must take care to avoid undercoverage. Suppose we plan to select 500 names from the city phone book, call their homes between noon and 4 p.m., and interview whoever answers, anticipating contacts with at least 200 people.

a) Why is it difficult to use a simple random sample here?

b) Describe a more convenient, but still random, sampling strategy.

c) What kinds of households are likely to be included in the eventual sample of opinion? Excluded?

d) Suppose, instead, that we continue calling each number, perhaps in the morning or evening, until an adult is contacted and interviewed. How does this improve the sampling design?

e) Random-digit dialing machines can generate the phone calls for us. How would this improve our design? Is anyone still excluded?

28. Cell phone survey. What about drawing a random sample only from cell phone exchanges? Discuss the advantages and disadvantages of such a sampling method compared with surveying randomly generated telephone numbers from non–cell phone exchanges. Do you think these advantages and disadvantages have changed over time? How do you expect they'll change in the future?

29. Arm length. How long is your arm compared with your hand size? Put your right thumb at your left shoulder bone, stretch your hand open wide, and extend your hand down your arm. Put your thumb at the place where your little finger is, and extend down the arm again. Repeat this a third time. Now your little finger will probably have reached the back of your left hand. If the fourth hand width goes past the end of your middle finger, turn your hand sideways and count finger widths to get there.

a) How many hand and finger widths is your arm?

b) Suppose you repeat your measurement 10 times and average your results. What parameter would this average estimate? What is the population?

c) Suppose you now collect arm lengths measured in this way from 9 friends and average these 10 measurements. What is the population now? What parameter would this average estimate?

d) Do you think these 10 arm lengths are likely to be representative of the population of arm lengths in your community? In the country? Why or why not?

30. Fuel economy. Occasionally, when I fill my car with gas, I figure out how many miles per gallon my car got. I wrote down those results after 6 fill-ups in the past few months. Overall, it appears my car gets 28.8 miles per gallon.

a) What statistic have I calculated?

b) What is the parameter I'm trying to estimate?

c) How might my results be biased?

d) When the Environmental Protection Agency (EPA) checks a car like mine to predict its fuel economy, what parameter is it trying to estimate?

31. Accounting. Between quarterly audits, a company likes to check on its accounting procedures to address any problems before they become serious. The accounting staff processes payments on about 120 orders each day. The next day, the supervisor rechecks 10 of the transactions to be sure they were processed properly.

a) Propose a sampling strategy for the supervisor.

b) How would you modify that strategy if the company makes both wholesale and retail sales, requiring different bookkeeping procedures?

32. Happy workers? A manufacturing company employs 14 project managers, 48 foremen, and 377 laborers. In an effort to keep informed about any possible sources of employee discontent, management wants to conduct job satisfaction interviews with a sample of employees every month.

a) Do you see any potential danger in the company's plan? Explain.

b) Propose a sampling strategy that uses a simple random sample.

c) Why do you think a simple random sample might not provide the representative opinion the company seeks?

d) Propose a better sampling strategy.

e) Listed below are the last names of the project managers. Use random numbers to select two people to be interviewed. Explain your method carefully.

Barrett	Bowman	Chen
DeLara	DeRoos	Grigorov
Maceli	Mulvaney	Pagliarulo
Rosica	Smithson	Tadros
Williams	Yamamoto	

33. **Quality control.** Sammy's Salsa, a small local company, produces 20 cases of salsa a day. Each case contains 12 jars and is imprinted with a code indicating the date and batch number. To help maintain consistency, at the end of each day, Sammy selects three jars of salsa, weighs the contents, and tastes the product. Help Sammy select the sample jars. Today's cases are coded 07N61 through 07N80.
    a) Carefully explain your sampling strategy.
    b) Show how to use random numbers to pick 3 jars.
    c) Did you use a simple random sample? Explain.

34. **A fish story.** Concerned about reports of discolored scales on fish caught downstream from a newly sited chemical plant, scientists set up a field station in a shoreline public park. For one week they asked fishermen there to bring any fish they caught to the field station for a brief inspection. At the end of the week, the scientists said that 18% of the 234 fish that were submitted for inspection displayed the discoloration. From this information, can the researchers estimate what proportion of fish in the river have discolored scales? Explain.

35. **Sampling methods.** Consider each of these situations. Do you think the proposed sampling method is appropriate? Explain.
    a) We want to know what percentage of local doctors accept Medicaid patients. We call the offices of 50 doctors randomly selected from local Yellow Page listings.
    b) We want to know what percentage of local businesses anticipate hiring additional employees in the upcoming month. We randomly select a page in the Yellow Pages and call every business listed there.

36. **More sampling methods.** Consider each of these situations. Do you think the proposed sampling method is appropriate? Explain.
    a) We want to know if there is neighborhood support to turn a vacant lot into a playground. We spend a Saturday afternoon going door-to-door in the neighborhood, asking people to sign a petition.
    b) We want to know if students at our college are satisfied with the selection of food available on campus. We go to the largest cafeteria and interview every 10th person in line.

JUST CHECKING
Answers

1. **a)** It can be hard to reach all members of a population, and it can take so long that circumstances change, affecting the responses. A well-designed sample is often a better choice.
    b) This sample is probably biased—students who didn't like the food at the cafeteria might not choose to eat there.
    c) No, only the sample size matters, not the fraction of the overall population.
    d) Students who frequent this Web site might be more enthusiastic about Statistics than the overall population of Statistics students. A large sample cannot compensate for bias.
    e) It's the population "parameter." "Statistics" describe samples.

2. **a)** systematic
    b) stratified
    c) simple
    d) cluster