- 12. Masking (also called blinding) can reduce bias due to demand characteristics—but it does not reduce social desirability bias. Making responses anonymous can reduce social desirability bias—but it does not reduce bias due to demand characteristics. Neither making participants blind nor making participants anonymous reduces random error.
- 13. Establishing internal consistency, discriminant validity, convergent validity, and content validity are all ways of building the case for a measure's construct validity.
- 14. With convergent validity, you are trying to show that you are measuring the right construct. Therefore, you show that your measure correlates with what it should

- correlate with—other measures or indicate of the construct you are trying to measure
- 15. With discriminant validity, you are trying show that you aren't measuring the wron construct. Therefore, you show that your measure does not correlate with what it should not correlate with—measures of different constructs.
- 16. Choosing a manipulation involves many of the same steps as choosing a measure.
- 17. Placebo treatments and unobtrusive measurement can reduce subject bias.
- 18. "Blind" (masked) procedures and standardization can reduce experimenter bias.
- 19. You can use manipulation checks to make a case for your manipulation's validity.

KEY TERMS

bias (p. 147)
blind masked (p. 152)
construct validity (p. 176)
content validity (p. 176)
convergent validity (p. 179)
demand characteristics (p. 156)
discriminant validity (p. 181)
environmental manipulation
(p. 189)
experimenter bias (p. 187)
Hawthorne effect (p. 186)

informed consent (p. 157)
instructional manipulation
 (p. 188)
internal consistency (p. 171)
interobserver (judge) agreement
 (p. 166)
interobserver (scorer) reliability
 coefficient (p. 166)
known-groups technique (p. 179)
manipulation check (p. 187)
observer bias (p. 150)

operational definitions (p. 144) placebo treatment (p. 186) random error (p. 147) reliable, reliability (p. 161) social desirability bias (p. 158) standardization (p. 154) stooges (confederates) (p. 190) subject bias (p. 155) test-retest reliability (p. 162) unobtrusive measurement (p. 157)

EXERCISES

- 1. Why is bias considered more serious than random error?
- 2. What are the two primary types of subject bias? What are the differences between these two types?
- 3. Suppose a "social intelligence" test in a popular magazine had high internal consistency. What would that mean? Why would you still want to see whether the test had discriminant validity? How would you do a study to determine whether the test had discriminant validity?
- 4. Given that IQ tests are not perfectly reliable, why would it be irresponsible to tell someone his or her score on an IQ test?

- 5. What is content validity? How does it differ from internal consistency? For what measures is it most important?
- 6. Swann and Rentfrow (2001) wanted to develop a test "that measures the extent to which people respond to others quickly and effusively." In their view, high scorers would tend to blurt out their thoughts to others immediately and low scorers would be slow to respond.
 - a. How would you use the known-groups technique to get evidence of your measure's construct validity?

- b. What measures would you correlate with your scale to make the case for your measure's discriminant validity? Why? In what range would the correlation coefficients between those measures and your measure have to be to provide evidence of discriminant validity? Why?
- c. To provide evidence of convergent validity, you could correlate scores on your measure with a behavior typical of people who blurt out their thoughts. What behavior would you choose? Why?
- 7. A researcher wants to measure "aggressive tendencies" and is trying to decide between a paper-and-pencil test of aggression and observing actual aggression.

a. What problems might there be with observing aggressive behavior?

- b. What would probably be the most serious threat to the validity of a paper-and-pencil test of aggression? What information about the test would suggest that the test is a good instrument?
- 8. Think of a construct you would like to measure.
 - a. Name that construct.
 - b. Define that construct.
 - c. Locate a published measure of that construct (if you are having trouble finding a published example, see Web Appendix B), and write down the reference for that source.
 - d. Develop a measure of that construct.

- e. What could you do to improve or evaluate your measure's reliability?
- f. If you had a year to try to validate your measure, how would you go about it? (Hint: Refer to the different kinds of validities discussed in this chapter.)
- g. How vulnerable is your measure to subject and observer bias? Why? Can you change your measure to make it more resistant to these threats?
- 9. What problems do you see with measuring "athletic ability" as 40-yard-dash speed? What steps would you take to improve this measure? (Hint: Think about solving the problems of bias and random error.)
- 10. Think of a factor that you would like to manipulate.
 - a. Define this factor as specifically as you
 - b. Find one example of this factor being manipulated in a published study (if you are having trouble finding a published example, see Web Appendix B). Write down the reference citation for that source.
 - c. Would you use an environmental or instructional manipulation? Why?
 - d. How would you manipulate that factor? Why?
 - e. How could you perform a manipulation check on the factor you want to manipulate? Would it be useful to perform a manipulation check? Why or why not?

WEB RESOURCES

- 1. Go to the Chapter 5 section of the book's student
 - a. Look over the concept map of the key terms.
 - b. Test yourself on the key terms.

- c. Take the Chapter 5 Practice Quiz.
- d. Do the interactive end-of-chapter exercises.
- 2. If you are ready to draft a method section, click on the "Method Section Tips" link.