

APPENDIX B - TEST YOURSELF KEY

CHAPTER 1

1. b
2. d
3. a
4. c
5. If we do not have conscious awareness of the conflict (i.e., if it is unconscious), then it would be difficult to support the existence of this conflict through direct observations of behavior. In other words, it is difficult to falsify this theory by directly observing behavior.
6. Parsimony (i.e., assuming that simple explanations are more likely to be correct) is useful in psychological studies because this assumption makes it easier to test explanations of behavior. It is easier to devise a study that can falsify a simple explanation than to devise a study to falsify a more complex explanation.
7. (a) applied, (b) basic, (c) basic, (d) applied, (e) basic, (f) applied, (g) basic, (h) basic
8. confirmation bias
9. observation
10. external validity
11. Observations (collecting data) are the best method of knowing in terms of objectivity and reducing bias in answers to questions we might have.
12. One good way to increase the external validity in this study would be to use a more realistic task, such as a topic content exam.

CHAPTER 2

1. (a) Results, (b) Method—Participants, (c) Introduction and Discussion, (d) Discussion, (e) Method—Procedure
2. Theory-driven hypotheses are made using predictions of a theory for a research study. Data-driven hypotheses are made by generalizing results from past studies to the results of a research study.

3. This is an inappropriate research question for psychological studies because it would be difficult to observe behaviors to answer this question. In other words, it is a question that cannot be answered by scientific methods because there are no clear behaviors that are seen as indicative of a soul.
4. A *peer-reviewed journal article* is typically an article written to describe a research study to inform other researchers what has previously been found in an area of psychology. *Peer reviewed* means that the article has been reviewed by experts in the area before publication to improve the article and make a recommendation about publication. An article in a popular magazine would likely provide a short summary of the study and may not accurately reflect the original article. Thus, magazine articles are considered secondary (rather than primary) sources.
5. A literature review is a review of previous research in an area to allow a researcher to conduct research that does not replicate what has previously been done and to allow researchers to make accurate predictions about a study's outcome.
6. (a) Behavior variable: conscientiousness; Causal variable: gender (men and women).
(b) Behavior variable: sleep quality; Causal variable: waking up at the same time.
(c) Behavior variable: violent behavior; Causal variable: ostracism. (d) Behavior variable: eyewitness memory; Causal variable: age (children and adults).
(e) Behavior variable: perceiving the size of an opening; Causal variable: using crutches.
7. PsycINFO will yield peer-reviewed articles about research. Google will yield other types of information (such as popular magazine articles) that are considered secondary sources that may not be reliable.
8. abstract
9. deductive
10. correlational
11. The primary difference between an empirical article and a book chapter or review article is that the empirical article will describe the details (purpose, method, results, and conclusions) of a single study (or set of studies), whereas book chapters and review articles summarize and organize a large set of studies by many researchers.
12. (a) correlational, (b) causal, (c) descriptive or correlational could be correct here, (d) correlational, (e) descriptive
13. Introduction: introduce topic, research question, related background studies, basics of the study design, and hypotheses; Method: describe participants, design, materials/apparatus/measures, and procedure of the study; Results: describe data in summary form and report statistical tests done to test hypotheses; Discussion: discuss how much support hypotheses received from the results, compare results with related studies, and state conclusions from the study

CHAPTER 3

1. As mentioned in the chapter, Burger (2009) conducted a modified replication of this study in conjunction with ABC News' *Primetime* show. To make the study ethical, and to have it approved by his IRB, he made the following changes: (a) Shock labels went to only 150 volts, at which point the confederate first objected verbally to the shocks. (b) Participants were prescreened to rule out individuals who might be more greatly affected by the stressful conditions of the study. (c) Right to withdraw from the study was emphasized (once verbally and twice in writing) to make this right more salient to participants. (d) The experimenter was a clinical psychologist who immediately stopped the study at the first sign of excessive stress from the participant.
2. The consent form should include the following: a statement that the participant is being asked to volunteer for research, that their participation is voluntary, and that they have the right to withdraw at any time without penalty; the purpose and the procedures of the study; the risks (including possible stress) and benefits of the study, and information about whom the participants can contact if they have questions or concerns about the study.
3. The debriefing process is especially important for research that involves deception because the goal is for participants to leave a study with as positive an impression of the research as possible. In addition, participants have the right to a thorough explanation of the study. Explaining why the deception was necessary is important in achieving these goals.
4. Answers will vary, but debriefing for this study should explain that the purpose of the study was to investigate natural physiological responses that occur to most individuals due to an invasion of personal space. It should also include information that the participant's data will be kept completely confidential, that the participant has the right to withdraw from the study at any time (especially if an informed consent process did not take place before the study), and information about whom the participants may contact if they have further questions or are feeling uncomfortable about participating in the study.
5. c
6. d
7. c
8. a
9. beneficence
10. plagiarism
11. (a) justice, (b) respect for persons, (c) beneficence, (d) respect for persons
12. The test of the hypothesis was not objective. In this case, the researcher is manipulating the data by collecting more data for the sole purpose of finding a result that matches their hypothesis. The data cannot be trusted in this study.

CHAPTER 4

1. c
2. a
3. b
4. d
5. There are lots of ways to test this. One example is to use a measurable work behavior (such as number of reports written in a week or number of clients served per day) as the operational definition of work productivity. The study could compare work productivity change (difference between work productivity before the study starts and work productivity during the study) for workers on different floors of the office (where one floor has the cappuccino machine and the other floor does not). Or work productivity could be measured for a week without the machine and then measured for a week with the machine and compared for the 2 weeks. Both designs involve a quasi-experiment. A true experiment would be difficult to conduct in the workplace.
6. (a) This is a causal hypothesis; thus, an experiment is the best design. (b) A manipulated independent variable that includes random assignment of participants to a ginkgo biloba group and a control group that receives a placebo should be included.
7. External; Internal
8. (a) archival data, (b) systematic observation, (c) survey/questionnaire, (d) naturalistic observation
9. negative
10. case study
11. (a) quasi-experimental (there will be TBI and control groups to compare and the groups cannot be randomly assigned ethically); (b) experimental (exercise and no daily exercise groups can be randomly assigned to determine the causal effect on anxiety); (c) experimental (music and no music groups can be randomly assigned to determine its causal effect on attention); (d) quasi-experimental (can group people based on their report of when they go to bed) or experimental (can randomly assign bedtimes to people to determine the causal effect)
12. IV1 = type of video game (social or matching), IV2 = time of test (before playing and after 2 weeks of play); DV = score on standardized test. Note: The study can also be conceptualized with only IV1 and DV = change in score on the test from time 1 to time 2.

CHAPTER 5

1. (a) ratio, (b) ordinal, (c) nominal, (d) ratio, (e) interval, (f) nominal, (g) ordinal
2. (a) ordinal, (b) ratio, (c) nominal, (d) interval
3. Independent variable = type of problem display, type independent variable, bivalent independent variable, manipulated between subjects.
4. With a between-subjects design, the primary concern is with group differences. Thus, one might be concerned about group differences for interest in math and ability (which can also affect interest). If group differences exist in math interest before the study, they cloud the test of the independent variable effect on the dependent variable.
5. The external validity can be improved with math problems and classroom situations that are more realistic. A field experiment can be done by designing different texts that present math problems in the different formats and by asking students to use them in the classroom. Tests can also be designed for classroom use with the different presentation formats.
6. regression toward the mean
7. double-blind
8. the Hawthorne effect
9. external
10. Attrition occurs when subjects drop out of a study before it is completed. The problem with attrition is that the subjects who drop out may be characteristically different from subjects who remain in the study. This can result in data that apply to only certain members of the group being studied, which can limit the conclusions a researcher can draw from the results of the study. See Chapter 6 for more discussion of bias in subject samples.
11. Counterbalancing involves presenting the independent variable levels (or multiple independent variable conditions) in different orders for different participants. This controls for order of conditions as a source of bias in the study because the different orders are spread equally across the participants.
12. One source of bias is how often the participants recall the story on their own in the time delays. Another source of bias could be encountering information that is related to the story in the delay times that could alter the participants' memory of the study. Other sources of bias are also possible.

CHAPTER 6

1. Probability samples are typically better representatives of the population they are drawn from because the probability of an individual being chosen is determined ahead of time and controlled through random sampling.
2. Many studies use convenience samples because it can be difficult to identify all the members of large populations to randomly select them.
3. There are several ways to do this, but the key is to use a convenience sample of students at the school in the same proportions as they exist at the school based on years in school and gender. These proportions will need to be known before the sample is selected. You will also need a recruitment technique (e.g., internet with email contact, surveys mailed to campus address) to describe how to administer your survey.
4. Internet samples tend to be more diverse than other samples, and they can be collected using fewer resources and in a shorter period of time.
5. d
6. b
7. b
8. population; sample
9. simple random
10. b
11. Cluster sample—this best matches a cluster sample because a group within the population is first identified (the people who have gone to the university's counseling center) and the people are randomly selected to be recruited for the study.
12. One advantage of an internet sample, such as MTurk, is that a researcher can reach a larger and more diverse set of participants than with a sample of college students. The sample can often be obtained more quickly this way as well. The disadvantages include bots completing the study or people paying less attention to the task than in person and needing resources for compensation (e.g., small payment) for the participants.
13. Using convenience samples (e.g., college students) has restricted the types of individuals studied in behavioral research based on age, socioeconomic status, and racial/ethnic diversity. This has led to lower external validity in many studies, as these samples do not represent large populations of people very well.

CHAPTER 7

1. (a) IV (independent variable) = video condition (view crime discussion video or not), DV (dependent variable) = recognition memory score. (b) This is a two-tailed test (no prediction is made about which group will score better), so for H_1 : in the population, the video condition affects recognition scores, for H_0 : in the population, the video condition does not affect recognition scores. (c) Reject the null hypothesis because $p < \alpha$: The independent variable (viewing the discussion of the crime with false information or not viewing the discussion) has an effect on memory accuracy. (d) If the null hypothesis is true and you reject it, you are making a Type I error.
2. (a) A one-tailed test should be used, because the researcher is predicting that people will be better than chance (> 50%) if they have poker abilities, but less than chance does not indicate poker abilities. (b) H_1 : in the population, better-than-average poker players will guess the type of card chosen at a rate higher than 50%, H_0 : in the population, better-than-average poker players will guess the type of card chosen at a rate lower than or equal to 50%. (c) $p > \alpha$, so the null hypothesis must be retained. The researcher cannot conclude that the poker players can predict another player's hand at better than chance accuracy. (d) This would be a Type II error, because the null hypothesis was retained when in reality it is false.
3. Type I errors occur when a researcher rejects a null hypothesis that is actually true. Type II errors occur when a researcher fails to reject a null hypothesis that is actually false. The probability of a Type I error is set by alpha.
4. A quasi-experiment: Two groups are compared, but no independent variable is manipulated.
5. one-tailed; two-tailed
6. reject
7. c
8. outlier
9. high
10. null
11. Both the mean and the mode are more affected by extreme scores in a distribution than the median. Thus, the median provides a better measure of central tendency for these distributions.
12. The larger the variability in a distribution, the less likely a small group (a sample) chosen from that population will be representative of that population. Thus, sampling error is estimated from the variability in a distribution of scores.

13. If a statistical test found a significant difference in the scores for freshmen and seniors, then there is good reason to believe that this difference exists in the population, as long as the study was well-designed. However, the difference does not fully test the authors' hypothesis because we would also need to know which group had the higher mean score.

CHAPTER 8

1. (a) Introduction and Discussion sections. (b) Results section. (c) Method section, Materials subsection. (d) Method section, Participants subsection. (e) Any of the major sections can contain a citation: Introduction, Method, Results, Discussion; the Abstract may also contain citations. (f) Method section, Procedure subsection.
2. (a) Only b and e are true
3. (a) Responses a and c are correct.
4. (a) Correct order: Title Page, Abstract, Introduction, Participants, Procedure, Results, Discussion, References.
5. (a) b
6. (a) An oral presentation provides an oral description of the major sections of a research report, typically with visual aids such as PowerPoint. There is often a time limit for oral presentations. Poster presentations provide a visual report of a research study, including each major section of the report. Authors are typically on site to provide oral summaries or more detailed information about the study the poster describes.
7. (a) Method
8. (a) Results
9. (a) Abstract
10. (a) an alphabetical list of all sources cited in a paper
11. (a) A thorough literature review should be conducted to ensure that you have read all relevant studies you can find to help explain the purpose of your study and justify your hypotheses. You will also want to ensure you remember all the details of the method to be able to correctly describe those details in your report. You should also have a clear understanding of the results of your data analyses and be prepared to describe the tests conducted and the statistical values obtained in those tests.
12. (a) A narrative citation is one in which the authors of a source are part of a sentence and the publication year is given in parentheses, such as, "Hancock et al. (2022) found results consistent with those reported here." A parenthetical citation is one in which the authors *and* the publication year are given in parentheses, such as "Previous studies have reported results with those found in the current study (e.g., Hancock et al., 2022)."

CHAPTER 9

1. (a) The design contains two groups of unmatched participants. Thus, the appropriate test is an independent samples t test to compare the means for the video discussion and no video discussion groups. (b) The null hypothesis tested in the analysis is that there is no mean difference between video discussion and no video discussion groups in the population. (c) SPSS output indicates that the video discussion group had mean recognition accuracy of 72.7% (SD = 5.23), and the no video discussion group had a mean recognition accuracy of 82.4% (SD = 7.52). The test indicates that the no video discussion group has a significantly higher mean recognition score than the video discussion group, $t(18) = -3.35, p = .004$. (d) Because the p value of .004 is less than alpha (.05), we can conclude that the null hypothesis is false and that there is a difference between video and no video groups in the population.
2. (a) The design contains a single sample of participants. The sample mean is compared with a known population mean when the participants have no poker ability (50% guessing accuracy). Thus, the appropriate test is a one-sample t test to compare the sample mean with the population mean of 50%. (b) The null hypothesis tested in the analysis is that there is no difference between the population mean of participants claiming to have poker abilities and the population mean of 50%. (c) SPSS output indicates that the sample mean is 50.4% (SD = 3.57). The test indicates that the sample mean does not significantly differ from 50%, $t(9) = 0.36, p = .73$. (d) Because the p value of .73 is greater than alpha (.05), we do not have evidence against the null hypothesis and must retain it.
3. (a) There are two independent variables in this experiment: ad type (youthfulness, attractiveness) and product type (car, energy drink). (b) This is a factorial design, and an ANOVA should be used to analyze the data. The main effect of each independent variable (ad type and product type) is tested to compare means for the levels (youthfulness vs. attractiveness, and car vs. energy drink) of that variable on its own, and an interaction effect between the two variables is also tested to determine if the effect of one independent variable depends on the level of the other variable. (c) Based on the p values reported in the output, the main effects of ad type, $F(1, 16) = 6.23, p = .02$, and product type, $F(1, 16) = 27.77, p < .001$, are both significant. However, the interaction between ad type and product type was not significant, $F(1, 16) = 0.08, p = .785$. (d) We can conclude that youthfulness ads ($M = 5.60, SD = 1.35$) result in higher desire-to-buy ratings than attractiveness ads ($M = 4.70, SD = 1.16$). We can also conclude that energy drink ads ($M = 6.10, SD = 0.74$) result in higher desire-to-buy ratings than car ads ($M = 4.20, SD = 1.03$).
4. d
5. c
6. a

CHAPTER 10

1. Most types of research questions can be answered with a survey if the behavior of interest can be observed through self-reports from the participants. However, most research that uses surveys addresses descriptive or predictive research questions.
2. open-ended; closed-ended
3. Test-retest reliability is good when scores on a survey show a strong positive correlation for participants who take the survey multiple times over a period of time. Split-half reliability is good when sets of scores on a survey show a strong positive correlation.
4. The question requires a forced choice between two alternatives, neither of which may be liked by the respondent. There are several ways to improve the question, but one way might be to use the following question and response choices:

If you were to have a pet, which animal would you prefer?

___ dog

___ cat

___ bird

___ reptile

___ other (please specify ____)

___ I do not like pets

5. Several aspects of the survey allowed the researchers to accurately predict the election outcome. For example, a random sample of likely voters instead of all Americans meant that they were sampling from people who were likely to cast a vote in the upcoming election. The large sample size kept the sampling error in the study low. The inclusion of cell phone users made the sample more representative, because some individuals in the population may not have landline phones.
6. d
7. c
8. a
9. e
10. b
11. Typically, closed-ended survey responses are analyzed with quantitative analyses (e.g., statistical tests, such as those described in Chapter 9), and open-ended survey responses are first analyzed in qualitative analyses to create a coding scheme for classifying the open-ended responses (which might later be followed up with some categorical quantitative analyses).

12. (a) Also include a response for longer times and for Never. (b) Need to reword question and response anchors to include responses for liking travel. (c) This is a double-barreled question and should be separated into two questions. (d) Need an option for None or 0. Grouping of numbers may not be optimal to capture variability in responses.

CHAPTER 11

1. This is a correlational study. No independent variable was manipulated, and other possible factors that could affect cancer rates were not controlled for. Thus, we cannot conclude that drinking green tea has a causal relationship with cancer rates. It is possible that it does, or it is possible that people who are less likely to get cancer (e.g., no family history) are more likely to drink green tea (i.e., low cancer probability causes one to drink green tea). Finally, it is possible that a third variable (e.g., maintaining a healthy lifestyle) causes one to both drink green tea and have low likelihood of getting cancer.
2. An experiment would need to be conducted where the researcher randomly assigns participants either to drink green tea every day or to not drink any green tea (but perhaps to drink a similar-looking drink that they are told is green tea but is actually just colored water) and then records who gets cancer in order to compare cancer rates across the two groups.
3. A descriptive research question is one where a researcher is interested in whether a behavior occurs and how often it occurs. In a correlational study, a descriptive research question might be whether a relationship exists between two variables and what type of relationship exists. A predictive research question is one in which a researcher examines whether one variable can be predicted from one or more variables.
4. (a) Weather as measured by temperature that day and barometric pressure that day, cognitive abilities, mood as measured by a questionnaire, activity level, and time spent outside that day. (b) Attitudes to math, math scores, and brain activity. (c) Predictor variables: family stress and functioning measured by surveys, outcome variable: expressed emotion measured by a survey. (d) Predictor variable: expectations about counseling measured by a survey before therapy; outcome variable: client-therapist working alliance measured by a survey after therapy.
5. (a) How often people exercise measured by how many times per week they exercise for 30 minutes or more at a time and happiness measured by the score on the quality-of-life scale. (b) More exercise may cause people to be happy, being happy could cause people to exercise more, or some third variable could cause both happiness and more exercise. (c) No variable is manipulated in this study; thus, no causal relationships are tested. One of the variables (it would be easiest to manipulate the amount of exercise variable) must be manipulated (e.g., participants randomly assigned to different conditions of the variable) to test a causal relationship.

6. positive
7. predictive
8. Correlational studies measure dependent variables from a single group of subjects to look for relationships between the measures, but do not separate subjects into groups for comparison as is done in quasi-experiments.
9. Type of relationship is a moderating variable because it changes the relationship that exists between the other variables.
10. If impulsivity is a mediating variable, this means that impulsivity might be the key predictor between age and time management skills (i.e., time management skills develop with age because impulsivity decreases with age).

CHAPTER 12

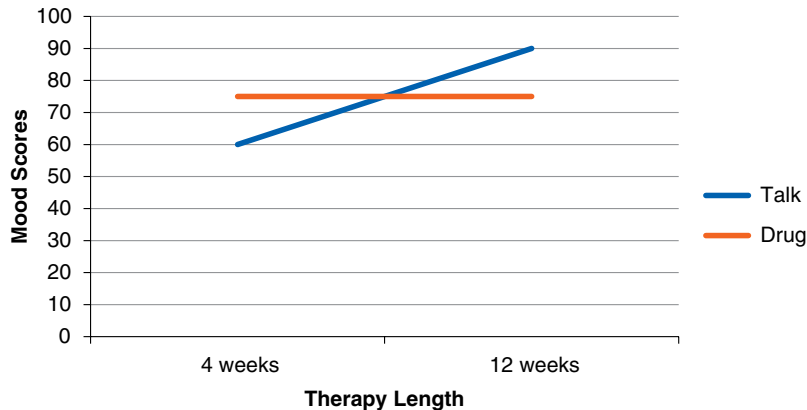
1. d
2. c
3. b
4. c
5. Manipulation of an independent variable and control of extraneous variables.
6. The Latin square allows fewer orders to be used than a full counterbalancing: useful for designs with more than three conditions. The number of orders in a Latin square matches the number of conditions in the experiment.
7. (a) The independent variable is the treatment for separation anxiety with levels of chew toy, our used objects, and confinement. (b) Within subjects. (c) Each treatment was used five times to control for length of time away and other possible variables that might change with our times away from home; the treatment orders were counterbalanced to control for order effects. (d) This experiment would likely be done by choosing a sample of dogs that represent the population and randomly assigning a portion of the sample to each of the treatments (e.g., 25 dogs to each treatment group). (See Chapter 6 for information about sampling techniques.) Confounding variables as described for the scenarios in the chapter are controlled for each treatment group (e.g., test all dogs for same length of visit away in same environment and at same time of day). This experiment uses a between-subjects design. The dependent variable is measured in the same manner as the chapter scenario with the operational definition specified.
8. (a) The independent variable is smile group: smile or no smile. It is manipulated between-subjects because the participants experience only one level of the

independent variable. (b) The main ethical issue in this study is that deception is used (i.e., the participants are misled about what the study is actually about and what behaviors are being observed). The researchers need to justify why the deception is necessary and then thoroughly debrief the subjects at the end of the experiment, so they understand why they were deceived and what the real purpose of the study was.

CHAPTER 13

1. a
2. b
3. a
4. b
5. No main effect of therapy type (each column has a mean of 75), main effect of therapy length (row means of 67.5 for 4 weeks vs. 82.5 for 12 weeks: longer has higher score), and interaction effect such that drug therapy has higher score for 4-week length, but talk therapy has higher score for 12-week length.
6. See the following sample graph:

FIGURE B1 ■ Example Graph Showing the Mean Mood Scores From Figure 13.A



7. (a) One independent variable is the treatment for separation anxiety with levels of chew toy and worn clothing. The other independent variable is length of outing with levels short and long. (b) Both were within-subjects. (c) Yes, the effectiveness of the treatment depended on the length of outing such that the chew toy was best for short outings and the worn clothing was best for long outings.

8. (a) Independent variable 1 = type of event; levels = false, true; independent variable 2 = type of false event; levels = criminal, noncriminal; dependent variables = ratings of anxiety, vividness, and confidence; also percentage of subjects who had false memory for event. (b) Independent variable 1 = type of trial; levels = near, far; quasi-independent variable 2 = age group; levels = adult, child; dependent variable = brain activity during the sessions. (c) Independent variable 1 = type of game; levels = inclusion, exclusion; independent variable 2 = perspective taken; levels = one of the players, none. (d) Independent variable = object locations; levels = organized by category, random organization; subject variable = age group with levels = 7-, 9-, and 11-year-olds, and adults; dependent variable = location errors (difference between location object is placed in during the test phase and the object's original location in the study phase).
9. (a) The independent variables are smile group: smile or no smile: It is manipulated between-subjects because the participants experience only one level of the independent variable and gender of thanking researcher—male or female—also between-subjects. (b) Yes, the effect of smile depended on the gender of the researcher such that smiles had an effect when the researcher was female, but not when the researcher was male.
10. (a) it is an experiment because there is a randomly assigned independent variable with age as a quasi-independent variable. (b) The independent variable is training group (practice vs. no practice are the levels), subject variable/quasi-independent variable is age at testing (levels of 4 and 6 years old). (c) Yes, these results show an interaction because the age difference *depends on* whether the children received the practice training.

CHAPTER 14

1. c
2. c
3. a
4. d
5. Manipulation of an independent variable and control for extraneous variables that can be done with random assignment.
6. The primary difference is that in a time series design, several pretest and posttest measures are recorded and the pattern of scores before and after a treatment or event is compared. In a simple pretest–posttest design, only one pretest and one posttest measure is recorded and the difference between these two scores indicates the effect of the treatment that occurred between the two measures.

7. Answers will vary, but the design should include a pretest and posttest on the same behavior, nonrandomized experimental and control groups, and a treatment or condition implemented for the experimental group between the pretest and posttest measures.
8. Designs will vary, but the study should include surveys of approval ratings each month or so for a period of several months before August 2011 and several months after August 2011.
9. The main sources of bias in a quasi-experiment are history effects (effects of a historic event on the results where the historic event is not the treatment of interest in the study), maturation effects (subjects in the study may develop or age over the course of study if it takes place over a longer period of time and this can influence or affect the results), and attrition or mortality (subjects drop out during the course of the study, possibly biasing the sample that completes the study).
10. This is an experiment because the participants are randomly assigned to the treatment/experimental (game related to cognitive aging) and control (game not related to cognitive aging) groups.
11. This is a quasi-experiment with a pretest-posttest design. It is simple comparing archival data from before the new law and after the new law with no control group.
12. It is an interrupted time series design. The *patterns* of scores are being compared before and after the pandemic (a natural/historic event) began.

CHAPTER 15

1. d
2. b
3. a
4. e
5. c
6. Answers will vary in terms of measures, but the longitudinal study will collect data from each child at multiple ages over time, and the cross-sectional study will compare age groups of different children at a single time.
7. The positive reinforcement treatment will vary, but in the A-B-A design, you would measure frequency of barking without the reinforcement for a set time (A—baseline), then measure frequency of barking during the same length of time with the reinforcement (B—treatment), then do another measure of the baseline behavior without the reinforcement to see if the reinforcement was actually reducing the barking (A).

