inferences about this variable (group: intervention vs. usual-care) is weakened by the lack of randomization.

Chapter 2

2-15. The scenario said nothing about randomly assigning participants to groups, so this study cannot be an experiment. All stimuli were offered to all participants, so there was no manipulation of an independent variable. Therefore, this study amounted to observational research with statistical replication (multiple participants being studied).

2-16. Stimulus is the predictor variable in this study.

2-17. Engagement duration is the criterion variable.

2-18. Participant's age is an extraneous variable. It might have an effect on the outcome variable (engagement duration) and is not being controlled because the study lacks randomization.

2-19. No. Causal conclusions cannot be drawn from a nonexperimental study.

2-20. Internal validity (the quality of inference about causal relationships between variables) is weak because this study is not an experiment.

2-21. External validity (quality of inference about generalization of results from the sample to the population) cannot be assessed without further information. There was no mention of random sampling from a population; in fact, the researchers used a convenience sample of residents in two nursing homes. We probably can generalize the results only to participants like the ones described in the article.

2-22. *N* is the sample size, which is 56.

2-23. The scenario said 44 out of the 56 residents were women, so the mode for the variable gender is "female." Because 35 of the 56 residents were widowed, the mode for marital status is "widowed." "High school or above" would be the mode for educational status because 47 of the 56 residents were in this category.

2-24. The range = 50. Notice how much more informative it is to know the minimum and maximum ages of residents in the study, compared with knowing "range = 50."

2-25. With the youngest participant being 61 and the oldest being 101, having a mean age of 87 could lead the reader to wonder whether the age of the oldest participant is inflating the mean. If the article had given the median, the reader could judge whether the higher age had affected the mean.

2-26. *SD* = 6.2 tells us something about the amount of variation in the cognitive functioning scores. It is the standard deviation computed on the unbiased variance, which is almost the average squared difference from the mean. By taking the square root of the unbiased variance, we get as close as we can to the average distance from the mean. Without information about the typical amount of variation in cognitive functioning scores, this statistic's magnitude is hard to judge.

2-27. No, the range would be 21 (that is, high score minus low score = 21 – 0 = 21).

2-28. This is an experimental study because there is random assignment to groups, manipulation of their experience (i.e., what was contained in the capsule they took every day) and statistical replication.

2-29. Treatment (glucosamine or placebo) is the independent variable.

2-30. RMDQ score is the dependent variable.

2-31. Usual therapies is an extraneous variable, which could interfere with the researchers' attempts to observe a relationship between treatment and the dependent variable. But random assignment should control this kind of extraneous variable.

2-32. The control group had more variation in its RMDQ scores because its *SD* = 4.5, which is a larger number than the treatment group's *SD* = 3.9. (The difference between 4.5 and 3.9 may be arbitrarily small.)

2-33. After a year it appeared that the control group had higher disability because its mean RMDQ = 5.5, compared with the treatment group's mean = 4.8. The researchers reported that the difference was not statistically remarkable.

2-34. Yes, we can make causal statements about this study because it is an experiment.

2-35. Mean SBP = 152.5, median SBP = 151, mean and median DBP = 70, mean and median HR = 71.

2-36. The middle reading for SBP is slightly lower than the arithmetic average. For the two other variables, the mean is the same as the middle score.

2-37. Subtracting the mean SBP = 152.5 from every score gives these results: -2.5, -0.5, -2.5, 11.5, 3.5, -9.5. Squaring the distances: 6.25, 0.25, 6.25, 132.25, 12.25, 90.25. Sum of squares =

numerator of the unbiased variance = 247.5. Denominator of the unbiased variance = $N - 1 = 6 - 1 = 5$. So the unbiased variance for SBP = 247.5 / 5 = 49.5.

2-38. Subtract the mean DBP = 70 from each score: -6, 5, 11, -5, 13, 8. Square each distance: 36, 25, 121, 25, 169, 64. Sum of squares = numerator of the unbiased variance = 440. Denominator of the unbiased variance = $N - 1 = 6 - 1 = 5$. So the unbiased variance for DBP = 440 / 5 = 88.

2-39. Subtract the mean HR from each score: 1, -1, -3, 3, -2, 2. Square the distances: 1, 1, 9, 9, 4, 4. Sum of squares = numerator of the unbiased variance = 28. Denominator of the unbiased variance = $N - 1 = 6 - 1 = 5$. So the unbiased variance for HR = 28 / 5 = 5.6.

2-40. DBP had more variability than SBP or HR, as evidenced by the unbiased variance being the largest for DBP. The HR readings showed the most consistency, because their unbiased variance was the smallest.

2-41. For SBP, $SD \approx 7.04$. For DBP, $SD \approx 9.38$. For HR, $SD \approx 2.37$. *SD* is reported instead of the unbiased variance because *SD* is in the original units of measure, whereas the unbiased variance is in squared units.

2-42. For SBP, $M = 152.5$ and $SD \approx 7.04$; for DBP, $M = 70$ and $SD \approx 9.38$; for HR, $M = 71$ and $SD \approx 2.37$. As one would expect for SBP, the central tendency of the readings is higher on the number line than the location of the middle of the DBP readings. But it appears there is more spread in the DBP readings (because the DBP readings' *SD* is larger than SBP's *SD*) and greater consistency in the SBP readings (because the SBP readings' *SD* is smaller than DBP's *SD*). Out of the three variables, HR has the smallest *SD*, so there is more consistency in those readings than in the BP measures.

2-43. The cases had been suffering from the pain for more than 10 years on average, but the middle score for number of months was 84, or 7 years.

2-44. The cases had a mean number of months since onset that is higher than the median. Similarly, they had a mean duration of RMMA episodes that is notably higher than the median. The means may have been pulled upward by a few extremely high numbers on each variable. Skewness statistics could tell us whether the distributions for these variables departed from symmetry. Based on the means and medians, we might expect to have seen positive skewness statistics.

2-45. The quoted material refers to a 5% trimmed mean. If 5% of scores are trimmed from each end of the distribution, the middle 90% of scores would remain to be averaged.

2-46. If a few mothers suffered from a great deal of postpartum depression, they may have extremely high scores on the maternity-blues scale. Extreme scores can pull the arithmetic mean toward those extremes. If no extreme scores were trimmed, the means might have been inflated by the experiences of a few women.

Chapter 3

3-7. Kind of stimulus is a nonnumeric, discrete categorical variable.

3-8. Engagement duration is a numeric or quantitative variable.

3-9. The scenario described the measurement of the amount of time that the residents engaged with each stimulus. A bar graph could be created with a separate bar for each stimulus. The mean amount of engagement time for each stimulus could be displayed.

3-10. A boxplot for each stimulus would be good for seeing how much variability there was among the participants' engagement times. Boxplots would show each condition's median, the amount of spread in the engagement times, and any outliers for each stimulus.

3-11. A bar graph could be used to show how many residents refused each stimulus. A separate bar would represent each stimulus, and the heights of the bars would show how many people refused the different stimuli.

3-12. A histogram and a boxplot could be made. The strength of a histogram is that we can see the entire distribution with quite a bit of detail at various numeric values. A disadvantage is that sometimes the graphing software can make a histogram with bars representing intervals, which can make it a little more difficult to understand the scores. An advantage of a boxplot is that we can identify outliers in an objective way. A disadvantage of a boxplot is that we lose quite a bit of detail about the scores in the distribution; we see the median, midrange, the range of each quarter of the data, and so forth, but not the actual scores. A boxplot also does not show any gaps that may exist in the middle of the distribution.

3-13. An outlier may stand out from the other scores in a histogram, but different people may look at a graph in different ways and disagree on whether a high score or a low score can be called an outlier. A boxplot provides a way to define whether an extreme score is an outlier.