CHAPTER 3

Bivariate Correlation

In Chapter 2, we looked at the various statistical procedures that researchers use when they want to describe single-variable data sets. We saw examples where data on two or more variables were summarized, but in each of those cases the data were summarized one variable at a time. Although there are occasions when these univariate techniques permit researchers to describe their data sets, most empirical investigations involve questions that call for descriptive techniques that simultaneously summarize data on more than one variable.

In this chapter, we will consider situations in which data on two variables have been collected and summarized, with interest residing in the relationship between the two variables. Not surprisingly, the statistical procedures that we will examine here are considered to be **bivariate** in nature. In a later chapter, we will consider techniques designed for situations wherein the researcher wishes to simultaneously summarize the relationships among three or more variables.

Three preliminary points are worth mentioning as I begin my effort to help you refine your skills at deciphering statistical summaries of bivariate data sets. First, the focus in this chapter will be on techniques that simply summarize the data. In other words, we are still dealing with statistical techniques that are fully descriptive in nature. Second, this chapter is similar to Chapter 2 in that we consider ways to summarize data that involve both picture and numerical indices. Finally, the material covered in the next chapter, Reliability and Validity, draws *heavily* on the information presented here. With these introductory points now behind us, let us turn to the central concept of this chapter, correlation.

The Key Concept behind Correlation: Relationship

Imagine that a researcher measures each of nine families with respect to two variables: average daily phone use (measured in minutes) and the number of teenagers within each family. The data for this imaginary group of families might turn out as follows:

Family	***Average Daily Phone Use (Minutes)***	***Number of Teenagers***
Abbott	75	2
Donatelli	100	3
Edwards	60	1
Franks	20	0
Kawasaki	70	2
Jones	120	4
Lopez	40	1
Meng	65	2
Smith	80	3

While it would be possible to look at each variable separately and say something about the central tendency, variability, and distributional shape of the nine scores (first for phone use, then for number of teenagers), the key concept of correlation requires that we look at the data on our two variables *simultaneously*. In doing this, we are trying to see (1) whether there is a **relationship** between the two sets of scores and (2) how strong or weak that relationship is, presuming that a relationship does in fact exist.

On a simple level, the basic question being dealt with by **correlation** can be answered in one of three possible ways. Within any bivariate data set, it *may* be the case that the high scores on the first variable tend to be paired with the high scores on the second variable (implying, of course, that low scores on the first variable tend to be paired with low scores on the second variable). I refer to this first possibility as the *high-high, low-low* case. The second possible answer to the basic correlational question represents the inverse of our first case. In other words, it *may* be the case that high scores on the first variable tend to be paired with low scores on the second variable (implying, of course, that low scores on the first variable tend to be paired with high scores on the second variable). My shorthand summary phrase for this second possibility is *high-low, low-high*. Finally, it is possible that little systematic tendency exists in the data at all. In other words, it *may* be the case that some of the high and low scores on the first variable are paired with high scores on the second variable while other high and low scores on the first variable are paired with low scores on the second variable. I refer to this third possibility simply by the three-word phrase *little systematic tendency*.

As a check on whether I have been clear in the previous paragraph, take another look at the hypothetical data presented earlier on the number of teenagers and amount of phone use within each of nine families. More specifically, indicate how that bivariate relationship should be labeled. Does it deserve the label *high-high, low-low*? Or the label *high-low, low-high*? Or the label *little systematic* *tendency*? If you haven't done so already, look again at the data presented and formulate your answer to this question.

To discern the nature of the relationship between phone use and number of teenagers, one must first identify each variable's high and low scores. The top three values for the phone use variable are 120, 100, and 80, while the lowest three values in this same column are 60, 40, and 20. Within the second column, the top three values are 4, 3, and 3; the three lowest values are 1, 1, and 0. After identifying each variable's high and low scores, the next (and final) step is to look at both columns of data simultaneously and see which of the three answers to the basic correlational question fits the data. For our hypothetical data set, we clearly have a *high-high, low-low* situation, with the three largest phone-use values being paired with the three largest number-of-teenagers values and the three lowest values in either column being paired with the low values in the other column.

The method I have used to find out what kind of relationship describes our hypothetical data set is instructive, I hope, for anyone not familiar with the core concept of correlation. That strategy, however, is not very sophisticated. Moreover, you won't have a chance to use it very often because researchers will almost always summarize their bivariate data sets by means of pictures, a single numerical index, a descriptive phrase, or some combination of these three reporting techniques. Let us now turn our attention to these three methods for summarizing the nature and strength of bivariate relationships.

Scatter Diagrams

Like histograms and bar graphs, a **scatter diagram** has a horizontal axis and a vertical axis. These axes are labeled to correspond to the two variables involved in the correlational analysis. The **abscissa** is marked off numerically so as to accommodate the obtained scores collected by the researcher on the variable represented by the horizontal axis; in a similar fashion, the **ordinate** is labeled so as to accommodate the obtained scores on the other variable. (With correlation, the decision as to which variable is put on which axis is fully arbitrary; the nature of the relationship between the two variables will be revealed regardless of how the two axes are labeled.) After the axes are set up, the next step involves placing a dot into the scatter diagram for each object that was measured, with the horizontal and vertical positioning of each dot dictated by the scores earned by that object on the two variables involved in the study.

In Excerpt 3.1, we see the raw data and a scatter diagram associated with a recent election in North Carolina. For each of the 28 counties in the First Congressional District, the researchers found out two things: (1) the percentage of black voters who were registered and (2) the percentage of votes cast for black candidates. Those data are presented as numbers in Table 3 and then as dots within Figure 1. The dot in the scatter diagram that is furthest to the right came from

EXCERPTS 3.2–3.3 • *Correlation Coefficients*

The total knowledge score was significantly correlated with the use of self-regulation skills ($r = .54$), but not with the other SCT variables.

Source: Suminski, R. R., and Petosa, R. (2006). Web-assisted instruction for changing social cognitive variables related to physical activity. *Journal of American College Health, 54*(4), p. 221.

However, time since injury was negatively associated ($r = -0.68$) with reaching one's initial goals.

Source: Wade, S. L., Michaud, L., and Brown, T. M. (2006). Putting the pieces together: Preliminary efficacy of a family problem-solving intervention for children with traumatic brain injury. *Journal of Head Trauma Rehabilitation, 21*(1), p. 63.

To help you learn how to interpret correlation coefficients, I have drawn a straight horizontal line to represent the continuum of possible values that will result from researchers putting data into a correlational formula:

−1.00 ———————— 0.00 ———————— +1.00

This correlational continuum will help you pin down the meaning of several adjectives that researchers use when talking about correlation coefficients and/or relationships: direct, high, indirect, inverse, low, moderate, negative, perfect, positive, strong, and weak.

First, consider the two halves of the correlational continuum. Any *r* that falls on the right side represents a **positive correlation;** this indicates a **direct relationship** between the two measured variables. (Earlier, I referred to such cases by the term *high-high, low-low.*) On the other hand, any result that ends up on the left side is a **negative correlation,** and this indicates an **indirect,** or **inverse, relationship** (i.e., *high-low, low-high*). If *r* were to land on either end of our correlation continuum, the term **perfect** could be used to describe the obtained correlation. The term **high** comes into play when *r* assumes a value close to either end (thus implying a **strong** relationship); conversely, the term **low** is used when *r* lands close to the middle of the continuum (thus implying a **weak** relationship). Not surprisingly, any *r* that ends up in the middle area of the left or right sides of our continuum will be called **moderate.**

In Excerpts 3.4 through 3.6, we see cases where researchers used adjectives to label their *r*s. In the first two of these excerpts, we see the concepts of weak and moderate being used to describe correlation coefficients. In the third excerpt, we see the concept of strong being used. Excerpt 3.6 is especially instructive because it shows that both positive and negative correlations are considered to be strong if they are far away from zero.

Before concluding our discussion of how to interpret correlation coefficients, I feel obligated to reiterate the point that when the issue of relationship is addressed,

EXCERPT 3.1 • *Raw Data and a Scatter Diagram*

TABLE 3 ***North Carolina First Congressional District Democratic Primary***

County	*% of Black Voters Registered*	*% of Votes Cast for Black Candidates*	*County*	*% of Black Voters Registered*	*% of Votes Cast for Black Candidates*
Wayne	8.4	11.7	Nash	49.7	60.9
Beaufort	22.6	34.2	Northampton	53.4	49.7
Columbus	29.4	71.3	Warren	55.1	70.0
Chowan	31.0	36.4	Vance	55.3	66.2
Perquimans	33.4	37.8	Pender	55.4	68.3
Bladen	35.8	58.7	Hertford	55.4	49.6
Greene	36.0	21.8	Pasquotank	55.7	63.8
Martin	39.5	31.1	Bertie	55.7	56.2
Pitt	42.1	33.5	Halifax	61.0	65.8
Washington	42.2	41.7	Lenoir	61.1	51.7
Duplin	43.3	51.5	Wilson	61.3	72.5
Craven	45.2	46.7	New Hanover	61.8	88.8
Gates	45.2	37.2	Cumberland	67.4	80.6
Jones	48.5	56.2	Edgecombe	72.4	79.0

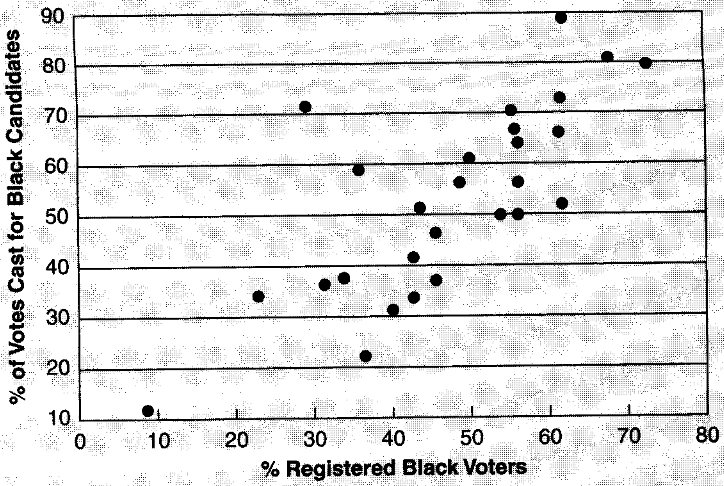

FIGURE 1 ***Black Voting in North Carolina First Congressional District***

Source: Clayton, D. M., and Stallings, A. M. (2000). Black women in Congress: Striking the balance. *Journal of Black Studies, 30*(4), pp. 593–594.

Edgecombe County where the two percentages were 72.4 and 79.0. Due to the way the axes were set up, that county's dot was positioned so as to correspond with 72.4 on the abscissa and 79.0 on the ordinate. All other dots in this scatter diagram were positioned in a similar fashion.

A scatter diagram reveals the relationship between two variables through the pattern that is formed by the full set of dots. To discern what pattern exists, I use a simple (though not completely foolproof) two-step method. First, I draw an imaginary perimeter line, or "fence" around the full set of data points—and in so doing, I try to achieve a tight fit. Second, I look at the shape produced by this perimeter line and examine its tilt and its thickness. Depending on these two characteristics of the data set's scatter, I arrive at an answer to the basic correlational question concerning the nature and strength of the relationship between the two variables.

Consider once again the scatter diagram shown in Excerpt 3.1. Our perimeter line produces a rough oval that is tilted from lower-left to upper-right. Tilts going in this direction imply a *high-high, low-low* relationship, whereas tilts going in the opposite direction, from upper-left to lower-right, imply a *high-low, low-high* relationship. (In cases where there is no discernible tilt to the shape produced by the perimeter line, there is little systematic tendency one way or the other.)

After establishing the tilt of the oval produced by our perimeter line, I then turn to the issue of the oval's thickness. If the oval is elongated and thin, then I conclude that there is a *strong* relationship between the two variables. On the other hand, if the oval is not too much longer than it is wide, then I conclude that a *weak* relationship exists. Considering one last time the scatter diagram in Excerpt 3.1, I conclude that the thickness of the oval produced by the perimeter line around the 34 dots falls between these two extremes; accordingly, I feel that the term *moderate* best describes the strength of the relationship that is visually displayed. Combining the notions of tilt and thickness, I feel that the scatter diagram in Excerpt 3.1 reveals a moderate *high-high, low-low* relationship between the two measured variables.

The Correlation Coefficient

Although a scatter diagram has the clear advantage of showing the scores for each measured object on the two variables of interest, many journals are reluctant to publish such pictures because they take up large amounts of space. For that reason, and also because the interpretation of a scatter diagram involves an element of subjectivity, numerical summaries of bivariate relationships appear in research reports far more frequently than do pictorial summaries. The numerical summary is called a **correlation coefficient.**

Symbolized as r, a correlation coefficient is normally reported as a decimal number somewhere between -1.00 and $+1.00$. In Excerpts 3.2 and 3.3, we see examples of correlation coefficients.

EXCERPTS 3.4–3.6 • *Use of Modifying Adjectives for the Term Correlation*

When the patients' scores on ASASFA were compared with their scores on the ADAS, the correlation ($r = 0.18$) was weak.

Source: McAdam, J. L., Stotts, N. A., Padilla, G., and Puntillo, A. (2005). Attitudes of critically ill Filipino patients and their families toward advance directives. *American Journal of Critical Care, 14*(1), p. 22.

The internalizing problem scale and the externalizing problem scale intercorrelated moderately (r = .37).

Source: Rönnlund, M., and Karlsson, E. (2006). The relation between dimensions of attachment and internalizing or externalizing problems during adolescence. *Journal of Genetic Psychology, 167*(1), p. 53.

Correlations between plasma caffeine and changes in performance were inconclusive, with the exception of a strong positive correlation with tackle sprint speed (r = 0.63) and a strong negative correlation with Drive 1 power (r = −0.80).

Source: Stuart, G. R., Hopkins, W. G., Cook, C., and Cairns, S. P. (2005). Multiple effects of caffeine on simulated high-intensity team-sport performance. *Medicine and Science in Sports and Exercise, 37*(11), p. 2001.

the central question being answered by r is: "To what extent are the high scores of one variable paired with the high scores of the other variable?" The term *high* in this question is considered separately for each variable. Hence, a strong positive correlation can exist even though the mean of the scores of one variable is substantially different from the mean of the scores on the other variable. As proof of this claim, consider again the data presented earlier on nine families who had varying numbers of teenagers and also varying amounts of phone use; the correlation between the two sets of scores turns out equal to +.96 despite the fact that the two means are quite different (2 versus 70). This example makes clear, I hope, the fact that a correlation does *not* deal with the question of whether two means are similar or different.[1]

The Correlation Matrix

When interest resides in the bivariate relationship between just two variables or among a small number of variables, researchers will typically present their rs within the text of their article. (This reporting strategy was used in Excerpts 3.2 through 3.6.)

[1]In many research studies, the focus is on the difference between means. Later, our discussion of t-test and F-tests will show how researchers compare means.

Research Navigator.com
Correlation matrix

When interest centers on the bivariate relationships among many variables, however, the resulting *r*s will often be summarized within a special table called a **correlation matrix.**

It should be noted that *several* bivariate correlations can be computed among a set of variables, even for relatively *small* sets of variables. With six variables, for example, 15 separate bivariate *r*s can be computed. With 10 variables, there will be 45 *r*s. In general, the number of bivariate correlations is equal to $k(k - 1)/2$, where k indicates the number of variables.

In Excerpt 3.7, we see a correlation matrix that summarizes the measured bivariate relationships among five variables. In the study associated with this excerpt, 344 college students in Taiwan responded to an inventory focused on the way people respond to stressful events. The inventory produced five scores, or factors, for each of the college students, with each of these scores corresponding to a mechanism used by people in the face of traumatic events. As you can see, this correlation matrix contains *r*s arranged in a triangle. Each *r* indicates the correlation between the two variables that label that *r*'s row and column. For example, the value of .37 is the correlation between Family Support and Religious/Spirituality.

EXCERPT 3.7 • *A Standard Correlation Matrix*

TABLE 2 *Intercorrelations among Factors of the CCS*

Factor	*1*	*2*	*3*	*4*	*5*
1. Acceptance, Reframing, and Striving	—				
2. Family Support	.33	—			
3. Religious/Spirituality	.13	.37	—		
4. Avoidance and Detachment	.15	.02	.05	—	
5. Private Emotional Outlets	.28	.34	.34	.29	—

Note: CSS = Collectivist Coping Styles

Source: Heppner, P. P., Heppner, M. J., Lee, D., Wang, Y., Park, H., and Wang, L. (2006). Development and validation of a collectivist coping styles inventory. *Journal of Counseling Psychology, 53*(1), p. 113.

Two things are noteworthy about the correlation matrix shown in Excerpt 3.7. First, when a row and a column refer to the same variable (as is the case with the top row and the left column, the second row and the second column, etc.), there is no correlation positioned at the intersection of that row and column. Instead, a dash has been put in each of those spots. This simply indicates that no one cares about the correlation of a variable with itself. (Such correlations, if they were computed, would be guaranteed to be equal to 1.00.)

The second thing to notice about the correlation matrix in Excerpt 3.7 is that there are no correlation coefficients above the diagonal formed by the dashes. If correlations did appear there, they would be a mirror image of the *r*s positioned below the diagonal. The value .33 would appear on the top row in the second column, .13 would appear on the top row in the third column, and so on. Such *r*s, if they were put into the correlation matrix, would be fully redundant with the *r*s that already are present; accordingly, they would add nothing.

In Excerpt 3.7, the correlation matrix was set up with the 10 bivariate correlation coefficients positioned below the diagonal of the dashes. At times, you will come across a correlation matrix in which (1) the values of the correlation coefficients are positioned above rather than below the diagonal or (2) each diagonal element has either 1.00 or nothing at all. Such alternative presentations should not cause you any difficulty, for they still will contain all possible bivariate correlations which are interpreted in the same way that we interpreted the *r*s in Excerpt 3.7.

Now consider the correlation matrix in Excerpt 3.8. On first glance, this one seems just like the one in Excerpt 3.7 because each of these correlation matrices has five rows and five columns of correlation coefficients. However, if you look closely at Excerpt 3.8, you will notice two things. First, the variable names used to label the rows are not exactly the same as those used to label the columns. The variable called Math, which labels the top row, does not appear as a label for any of the columns. Similarly, the variable called Spelling, which labels the right column, does not appear as a label for any of the rows. Second, the diagonal is filled with correlation coefficients.

If the correlation matrix in Excerpt 3.8 had been set up with a bottom row called Spelling and a left column called Math, it would have resembled the first correlation matrix we considered, except this one would have nothing at all in the

EXCERPT 3.8 • *A Correlation Matrix with One Row and One Column Deleted*

TABLE 3 ***Correlations among Self-Perception Items***

	Science	***Social Studies***	***Reading***	***English***	***Spelling***
Math	.306	.297	.217	.318	.265
Science		.429	.252	.337	.241
Social Studies			.331	.386	.299
Reading				.483	.369
English					.429

Source: Swiatek, M. A. (2005). Gifted students' self-perceptions of ability in specific subject domains: Factor structure and relationship with above 3-level test scores. *Roeper Review, 27*(2), p. 106.

diagonal. However, nothing would be gained by the addition of those new rows and columns. That's because neither the new bottom row nor the new left column would have any entries. (If you look again at Excerpt 3.7, you will discover that one row and one column did not contain any correlation coefficients of interest.)

Occasionally, researchers will set up their correlation matrices like the one in Excerpt 3.8. By deleting one empty row and one column, a little space is saved. Knowing that this is sometimes done, you must be careful when trying to figure out how many variables were involved; simply counting the number of rows (or columns) may cause you to end up one variable short.

Excerpt 3.9 illustrates how two correlation matrices can be combined into one table. In the study associated with this table, a personality inventory that yields five main scores (for emotional stability, extroversion, openness, agreeableness, and conscientiousness) was administered to a group of 162 undergraduates twice, once in a paper-and-pencil format and once over the Internet. After collecting the data, the researchers computed bivariate correlation coefficients among the personality dimensions. They did this twice, once using the data from the paper-and-pencil version and then again using data from the online version. Using the note beneath the correlation matrix as a guide, we can look to see if the bivariate correlation between any two dimensions was influenced very much by the way the personality inventory was administered. For example, the correlation between openness and extroversion was .54 in the Internet version compared to .51 in the paper-and-pencil version.

EXCERPT 3.9 • *Two Correlation Matrices Combined into One Table*

TABLE 2 ***Intercorrelations among the Big Five Personality Dimensions in the Paper-and-Pencil and Internet-Based Versions***

	ES	*EX*	*OP*	*AG*	*CO*
ES	—	.32	.15	.17	−.08
EX	.27	—	.54	.00	.03
OP	.11	.51	—	.00	.08
AG	.20	−.06	−.08	—	−.08
CO	−.13	.01	.12	−.10	—

Note: Correlations below the diagonal correspond to the paper-and-pencil version and the correlations above the diagonal correspond to the Internet-based version. $N = 162$.

Source: Salgado, J. F., and Moscoso, S. (2003). Internet-based personality testing: Equivalence of measures and assessees' perceptions and reactions. *International Journal of Selection & Assessment, 11*(2/3), p. 199.

Now consider Excerpt 3.10. This correlation matrix is different in two ways from the others we have examined. First, the mean and standard deviation are

EXCERPT 3.10 • *A Correlation Matrix with Means and Standard Deviations*

TABLE 3 ***Descriptive Statistics and Intercorrelations for Study Variables***

Variable	*M*	*SD*	*1*	*2*	*3*	*4*	*5*
1. Injuries	13.91	4.25	—				
2. Safety Events	22.83	7.33	.68	—			
3. Safety Climate	34.00	5.74	−.42	−.47	—		
4. Safety Consciousness	3.96	0.69	−.26	−.29	.63	—	
5. Passive Leadership	2.19	0.99	.33	.41	−.57	−.45	—
6. Transformational Leadership	3.08	0.90	−.24	−.22	.56	.41	−.48

Source: Kelloway, E. K., Mullen, J., and Francis, L. (2006). Divergent effects of transformational and passive leadership on employee safety. *Journal of Occupational Health Psychology, 11*(1), p. 8.

presented for each of the six study variables. Second, there are six rows but only five columns in the correlation matrix part of the table. All bivariate *r*s are presented, however. The researchers did not deprive us of anything by failing to include a sixth column (labeled 6) in the correlation matrix. Had they included it, the only thing in it would have been a dash.

Different Kinds of Correlational Procedures

In this section, we take a brief look at several different correlational procedures that have been developed. As you will see, all of these techniques are similar in that they are designed for the case in which data have been collected on two variables.[2] These bivariate correlational techniques differ, however, in the nature of the two variables. In light of this important difference, you need to learn a few things about how variables differ.

The first important distinction that needs to be made in our discussion of variables is between quantitative and qualitative characteristics. With a **quantitative variable,** the targets of the measuring process vary as to how much of the characteristic is possessed. In contrast, a **qualitative variable** comes into play when the things being measured vary from one another in terms of the categorical group to

[2]Some authors use the term **zero-order correlation** when referring to bivariate correlations. They do this to distinguish this simplest kind of correlation—that involves data on just two variables—from other kinds of correlations that involve data on three or more variables (such as partial correlations, multiple correlations, and canonical correlations).

which they belong relative to the characteristic of interest. Thus if we focus our attention on people's heights, we have a quantitative variable (because some people possess more "tallness" than others). If, on the other hand, we focus our attention on people's favorite national park, we would be dealing with a qualitative variable (because people simply fall into categories based on which park they like best).

From the standpoint of correlation, quantitative variables can manifest themselves in one of two ways in the data a researcher collects. Possibly, the only thing the researcher will want to do is order individuals (or animals, or objects, or whatever) from the one possessing the greatest amount of the relevant characteristic to the one possessing the least. The numbers used to indicate ordered position normally are assigned such that 1 goes to the person with the greatest amount of the characteristic, 2 goes to the person with the second greatest amount, and so on. Such numbers are called **ranks** and are said to represent an **ordinal** scale of measurement. A researcher's data would also be ordinal in nature if each person or thing being measured is put into one of several ordered categories, with everyone who falls into the same category given the same score. (For example, the numbers 1, 2, 3, and 4 could be used to represent freshmen, sophomores, juniors, and seniors.)

With a second kind of quantitative variable, measurements are more precise. Here, the score associated with each person supposedly reveals how much of the characteristic of interest is possessed by that individual—and it does this without regard for the standing of any other measured person. Whereas ranks constitute data that provide relative comparisons, this second (and more precise) way of dealing with quantitative variables provide absolute comparisons. In this book, we will use the term **raw score** to refer to any piece of data that provides an absolute (rather than relative) assessment of one's standing on a quantitative variable.[3]

Qualitative variables come in two main varieties. If the subgroups into which people are classified truly have no quantitative connection with each other, then the variable corresponding to those subgroups is said to be **nominal** in nature. Your favorite academic subject, the brand of jelly you most recently used, and your state of residence exemplify this kind of variable. If there are only two categories associated with the qualitative variable, then the variable of interest is said to be **dichotomous** in nature. A dichotomous variable actually can be viewed as a special case of the nominal situation, with examples being "course outcome" in courses where the only grades are pass and fail (or credit and no credit), gender, party affiliation during primary elections, and graduation status following four years of college.

In Excerpts 3.11 through 3.14, we see examples of different kinds of variables. The first two of these excerpts illustrate the two kinds of quantitative variables we have discussed: ranks and raw scores. The last two of these excerpts

[3]Whereas most statisticians draw a distinction between interval and ratio measurement scales and between discrete and continuous variables, readers of journal articles do not need to understand the technical differences between these terms in order to decipher research reports.

EXCERPTS 3.11–3.14 • *Different Kinds of Data*

The third step of the interview asked participants to rank, from best to worst, a set of randomly presented cards representing four health states: (a) your health, (b) mild stuttering, (c) moderate stuttering, and (d) severe stuttering.

Source: Bramlett, R. E., Bothe, A. K., and Franic, D. M. (2006). Using preference-based measures to assess quality of life in stuttering. *Journal of Speech, Language, and Hearing Research, 49*(2) p. 386.

We obtained sleep data by using Actiwatch (Mini Mitter, Bend, Oregon) activity monitors, which detected each participant's movement. . . . Reported as a percentage, sleep efficiency is defined as the total sleep time divided by the total time in bed multiplied by 100.

Source: Arora, V., Dunphy, C., Chang, V. Y., Ahmad, F., Humphrey, H. J., and Meltzer, D. (2006). The effects of on-duty napping on intern sleep time and fatigue. *Annals of Internal Medicine, 144*(11), p. 793.

The predictor variable race was a polytomous variable with four levels, as defined by the RSA in the national database (White, Black, American Indian or Alaskan Native, and Asian or Pacific Islander).

Source: Rogers, J. B., Bishop, M., and Crystal, R. M. (2005). Predicting rehabilitation outcome for supplemental security income and social security disability income recipients: Implications for consideration with the ticket to work program. *Journal of Rehabilitation, 71*(3), p. 8.

The type of school each student attended (public or parochial) was assessed, and a dichotomous score for school type was created, with 1 representing public schools.

Source: Nichols, T. R., Graber, J. A., Brooks-Gunn, J., and Botvin, G. J. (2006). Ways to say no: Refusal skill strategies among urban adolescents. *American Journal of Health Behavior, 30*(3), p. 230.

exemplify qualitative variables (the first being a four-category nominal variable, the second being a dichotomous variable).

Researchers frequently will derive a raw score for each individual being studied by combining that individual's responses to the separate questions in a test or survey. As Excerpts 3.15 and 3.16 show, the separate items can each be ordinal or even dichotomous in nature, and yet the sum of those item scores is looked upon as being what I have called a raw score. Although theoretical statistical authorities argue back and forth as to whether it is prudent to generate raw scores by combining ordinal or dichotomous data, doing so is an extremely common practice among applied researchers.

EXCERPTS 3.15–3.16 • *Combining Ordinal or Dichotomous Data to Get Raw Scores*

The resultant 31-item SES was scored on a 1- to 4-point Likert scale (1 = *never,* 2 = *seldom,* 3 = *fairly often,* 4 = *frequently*). A student's total score was determined by summing the responses on the 31 items, for a scale range of 31 to 124.

Source: Phillips, K. A., and Barrow, L. H. (2006). Investigating high school students' science experiences and mechanics understanding. *School Science & Mathematics, 106*(4), p. 204.

The VR-based pre-test and VR-based post-test that were employed in this study were computer based. Each test consisted of 15 questions and aimed to assess the learners' understanding of traffic rules and traffic signs. . . . For each question, participants received a score of either 1 (correct answer) or 0 (incorrect answer), and a total score ranging from 0 to 15.

Source: Chwen, J. C., Seong, C. T., and Wan Moh, F. W. I. (2005). Are learning styles relevant to virtual reality? *Journal of Research on Technology in Education, 38*(2), p. 128.

One final kind of variable needs to be briefly mentioned. Sometimes a researcher will begin with a quantitative variable but then classify individuals into two categories on the basis of how much of the characteristic of interest is possessed. For example, a researcher conceivably could measure people in terms of the quantitative variable of height, place each individual into a tall or short category, and then disregard the initial measurements of height (that took the form of ranks or raw scores). Whenever this is done, the researcher transforms quantitative data into a two-category qualitative state. The term **artificial dichotomy** is used to describe the final data set. An example of this kind of data conversion appears in Excerpt 3.17.

EXCERPT 3.17 • *Creating an Artificial Dichotomy*

Maternal reports of annual household income were used as a marker for socioeconomic disadvantage. Respondents were given a list of income ranges (e.g., $21,000–$23,999) from "no income" to "$150,000 and over" and were asked to choose the range in which their annual household income fell. Annual household income was made into a dichotomous variable for model-fitting analyses, using $24,000 as a cut point. This created a "high" income group ($24,000 and over) and a "low" income group ($0–$23,999).

Source: Cronk, N. J., Slutske, W. S., Madden, P. A. F., Bucholz, K. K., and Heath, A. C. (2004). Risk for separation anxiety disorder among girls: Paternal absence, socioeconomic disadvantage, and genetic vulnerability. *Journal of Abnormal Psychology, 113*(2), p. 240.

Pearson's Product-Moment Correlation

Research Navigator.c@m

Pearson's product-moment correlation

The most frequently used bivariate correlational procedure is called **Pearson's product-moment correlation.** It is designed for the situation in which (1) each of the two variables is quantitative in nature and (2) each variable is measured so as to produce raw scores. The scatter diagram presented earlier in Excerpt 3.1 provides a good example of the kind of bivariate situation that is dealt with by means of Pearson's technique.

Excerpts 3.18, 3.19, and 3.20 illustrate the use of this extremely popular bivariate correlational technique. Note, in the second of these excerpts, that the label "Pearson" is used by itself without the follow-up phrase "product-moment." In the third excerpt, note that only the symbol *r* is presented, and there is no adjective such as Pearson's, Pearson's product-moment, or product-moment. (In cases like this, where the symbol *r* stands by itself without a clarifying label, it's a good bet you're looking at a Pearson product-moment correlation coefficient.)

EXCERPTS 3.18–3.20 • *Pearson's Product-Moment Correlation*

The Pearson product-moment correlation coefficients of the participants' MTEBI and STEBI scores and math anxiety scores were calculated to explain the possible relationships between these variables.

Source: Bursal, M., and Paznokas, L. (2006). Mathematics anxiety and preservice elementary teachers' confidence to teach mathematics and science. *School Science & Mathematics, 106*(4), p. 175.

First, Pearson correlation coefficients were calculated for the physical performance measures (SMW, TUG, and STR) and all of the psychosocial and mechanical variables.

Source: Maly, M. R., Costigan, P. A., and Olney, S. J. (2005). Contribution of psychosocial and mechanical variables to physical performance measures in knee osteoarthritis. *Physical Therapy, 85*(12), p. 1323.

BD [body dissatisfaction] and DT [drive for thinness] scores correlated ($r = .72$ in females and $r = .65$ in males).

Source: Keski-Rahkonen, A., Bulik, C. M., Neale, B. M., Rose, R. J., Rissanen, A., and Kaprio, J. (2005). Body dissatisfaction and drive for thinness in young adult twins. *International Journal of Eating Disorders, 37*(3), p. 191.

Spearman's Rho and Kendall's Tau

The second most popular bivariate correlational technique is called **Spearman's rho.** This kind of correlation is similar to the one we just discussed (Pearson's) in that it is appropriate for the situation in which both variables are quantitative in

Research Navigator.com

Rank-order correlation

nature. With Spearman's technique, however, each of the two variables is measured in such a way as to produce ranks. This correlational technique often goes by the name **rank-order correlation** (instead of Spearman's rho). The resulting correlation coefficient, if symbolized, is usually referred to as r_s or ρ.

In Excerpt 3.21, we see two sets of ranks. This table comes from an article in which the researchers wanted to see what kinds of medical problems plagued psychiatry patients who had been seen in emergency departments of hospitals. There were two groups of psychiatric patients: (1) those whose primary problem was psychiatric in nature and (2) those whose primary problem was nonpsychiatric. What the researchers ranked within each of these two groups was the co-occurrence of various medical problems. If we correlate the two sets of ranks in Excerpt 3.21, using Spearman's rho, it turns out that $r_s = .89$. (If you scan the two sets of ranks in the excerpt, you should be able to see a *high-high, low-low* relationship.)

EXCERPT 3.21 • *Two Sets of Ranks*

TABLE 2 ***Percentage of Patients in Each Medical Category with Either a Primary or Secondary Psychiatric Diagnosis***

	Patients with a psychiatric diagnosis as primary			*Patients with a psychiatric diagnosis as secondary*		
Medical category	*Rank*	*%*[a]	*n*	*Rank*	*%*[b]	*n*
Circulatory	1	12.61	104	1.5	13.71	71
Endocrine	2	7.64	63	3	9.65	50
Respiratory	3	3.27	27	4	9.46	49
Digestive	4.5	2.42	20	1.5	13.71	71
Musculoskeletal	4.5	2.42	20	5	7.92	41
Nervous system	6	1.58	13	7.5	4.63	24
Infections	7	1.33	11	6	6.18	32
Blood	8	1.09	9	10.5	2.12	11
Genitourinary	9.5	0.49	4	9	4.05	21
Skin	9.5	0.49	4	10.5	2.12	11
Congenital	11.5	0.12	1	7.5	4.63	24
Neoplasms	11.5	0.12	1	13	0.39	2
Pregnancy	13	0.00	0	12	1.35	7

[a]Percentage of patients with a primary psychiatric diagnosis and no secondary psychiatric diagnosis.

[b]Percentage of patients with a secondary psychiatric diagnosis and no primary psychiatric diagnosis.

Source: Kunen, S., Niederhauser, R., Smoth, P. O., Morris, J. A., and Marx, B. D. (2005). Race disparities in psychiatric rates in emergency departments. *Journal of Consulting and Clinical Psychology, 73*(1), p. 121. (Modified slightly for presentation here.)

Only rarely will a researcher display the actual ranks utilized to compute Spearman's rho. Most of the time, the only information you will be given will be (1) the specification of the two variables being correlated and (2) the resulting correlation coefficient. Excerpts 3.22 and 3.23, therefore, are more typical of what you will see in published journal articles than is the material in Excerpt 3.21.

EXCERPTS 3.22–3.23 • *Spearman's Rank-Order Correlation*

Spearman's rank correlation coefficients were used to determine the correlation between the clinic-based gait velocity over 10 m and the overall time it took to complete the community walking course.

Source: Taylor, D., Stretton, C. M., Mudge, S., and Garrett, N. (2006). Does clinic-measured gait speed differ from gait speed measured in the community in people with stroke? *Clinical Rehabilitation, 20*(5), p. 440.

Spearman's correlation coefficient (r_s) between the change in CPH42 Profile score with the change in pain intensity at week 3 after the beginning of physiotherapy was 0.45.

Source: Chiu, T. T. W., Tai-Hing, L., and Hedley, A. J. (2005). Psychometric properties of a generic health measure in patients with neck pain. *Clinical Rehabilitation, 19*(5), p. 508.

Research Navigator.c⊕m
Kendall's tau

Kendall's tau is very similar to Spearman's rho in that both of these bivariate correlational techniques are designed for the case in which each of two quantitative variables is measured in such a way as to produce data in the form of ranks. The difference between rho and tau is related to the issue of ties. To illustrate what we mean, suppose six students took a short exam and earned these scores: 10, 9, 7, 7, 5, and 3. These raw scores, when converted to ranks, become 1, 2, 3.5, 3.5, 5, and 6, where the top score of 10 receives a rank of 1, the next-best score (9) receives a rank of 2, and so on. The third- and fourth-best scores tied with a score of 7, and the rank given to each of these individuals is equal to the mean of the separate ranks that they would have received if they had not tied. (If the two 7s had been 8 and 6, the separate ranks would have been 3 and 4, respectively; the mean of 3 and 4 is 3.5, and this rank is given to each of the persons who actually earned a 7.)

Kendall's tau is simply a bivariate correlational procedure that does a better job of dealing with tied ranks than does Spearman's rho. For the two sets of ranks shown in Excerpt 3.21, Kendall's tau turns out equal to .75, which is smaller than the Spearman rho value of .89 derived from the same set of ranks. Because there were tied ranks for some of the medical categories within each set of patients, many statisticians would consider tau to be the more accurate of these two correlation coefficients. In Excerpt 3.24, we see a case where Kendall's tau was used.

EXCERPT 3.24 • *Kendall's Tau*

The anaesthetists' VAS scores for elderly patients' correlated well with their anxiety for anaesthesia and surgery (Kendall's $\tau = 0.647$ and 0.524) and the surgeons made moderate estimation of patients' surgery-related anxiety on those undergoing major surgery (Kendall's $\tau = 0.480$).

Source: Fekrat, F., Sahin, A., Yazici, K. M., and Aypar, U. (2006). Anaesthetists' and surgeons' estimation of preoperative anxiety by patients submitted for elective surgery in a university hospital. *European Journal of Anaesthesiology, 23*(3), p. 230.

Point Biserial and Biserial Correlations

Research Navigator.com
Point biserial

Sometimes a researcher will correlate two variables that are measured so as to produce a set of raw scores for one variable and a set of 0s and 1s for the other (dichotomous) variable. For example, a researcher might want to see if a relationship exists between the height of basketball players and whether they score any points in a game. For this kind of bivariate situation, a correlational technique called **point biserial** has been designed. The resulting correlation coefficient is usually symbolized as r_{pb}.

If a researcher has data on two variables where one variable's data are in the form of raw scores while the other variable's data represent an artificial dichotomy, then the relationship between the two variables will be assessed by means of a technique called **biserial correlation.** Returning to our basketball example, suppose a researcher wanted to correlate height with scoring productivity, with the second of these variables dealt with by checking to see whether each player's average is less than 10 points or some value in the double digits. Here, scoring productivity is measured by imposing an artificial dichotomy on a set of raw scores. Accordingly, the biserial techniques would be used to assess the nature and strength of the relationship between the two variables. This kind of bivariate correlation is usually symbolized by r_{bis}.

In Excerpt 3.25, we see a case where the point-biserial correlation was used in a published research article. In this study, the researchers collected data to see if

EXCERPT 3.25 • *Point Biserial Correlation*

Treating the household as the unit of analysis and employing the age of the household head to identify the household's position in the life cycle, we indeed find home-ownership to increase with age. . . . Stated more formally, the point-biserial correlation coefficient—used to estimate the degree of relationship between a binary and a continuous variable—between home-ownership and age is positive with a value of 0.29.

Source: Başlevent, C., and Dayioglu, M. (2005). The effect of squatter housing on income distribution in urban Turkey. *Urban Studies, 42*(1), p. 39.

home ownership and age were correlated. As indicated in this excerpt, the variable of home ownership was binary (i.e., dichotomous) in nature because each household head either did or didn't own the home. Because the second variable, age, was a raw score variable, a point-biserial correlation was utilized to assess the relationship between the two variables.

Phi and Tetrachoric Correlations

If both of a researcher's variables are dichotomous in nature, then the relationship between the two variables will be assessed by means of a correlational technique called **phi** (if each variable represents a true dichotomy) or a technique called **tetrachoric correlation** (if both variables represent artificial dichotomies). An example calling for the first of these situations would involve, among high school students, the variables of gender and car ownership; since each variable represents a true dichotomy, the correlation between gender (male/female) and car ownership (yes/no) would be accomplished using phi. For an example of a place where tetrachoric correlation would be appropriate, imagine that we measure each of several persons in terms of height (with people classified as tall or short depending on whether or not they measure over 5'8") and weight (with people classified as "OK" or "not OK" depending on whether or not they are within 10 pounds of their ideal weight). Here, both height and weight are forced into being dichotomies.

Excerpt 3.26 illustrates the use of phi. This excerpt shows nicely how the two variables involved in a correlation can each represent a true dichotomy.

EXCERPT 3.26 • *Phi*

We computed a phi coefficient to estimate the degree of relationship between the intervention score and posttest score. [The first of these scores was a 1 or 0 depending on whether the single intervention question was answered correctly or incorrectly; likewise, the second of these scores was a 1 or 0 depending on whether the single posttest question was answered correctly or incorrectly.]

Source: Fernandez-Berrocal, P., and Santamaria, C. (2006). Mental models in social interaction. *Journal of Experimental Education, 74*(3), p. 235.

Cramer's V

If a researcher has collected bivariate data on two variables where each variable is nominal in nature, the relationship between the two variables can be measured by means of a correlational technique called **Cramer's *V*.** In Excerpt 3.27, we see a case where Cramer's *V* was used. In this study focused on a group of teenagers, two of the variables of interest were abstinence and parental marital status. As indicated in the excerpt, Cramer's *V* measured the strength of the relationship between these variables. This correlational technique yields coefficients that must lie somewhere between 0 and 1.

EXCERPT 3.27 • *Cramer's V*

More of sexually abstinent teens (50.3%) reported that their parents are married, compared to 32.4% of their sexually active peers (Cramer's $V = .18$).

Source: Vélez-Pastrana, M. C., Gonzádez-Rodríguez, R. A., and Borges-Hernández, A. (2005). Family functioning and early onset of sexual intercourse in Latino adolescents. *Adolescence, 40*(160), p. 787.

Warnings about Correlation

At this point, you may be tempted to consider yourself a semiexpert when it comes to deciphering discussions about correlation. You now know what a scatter diagram is, you have looked at the correlational continuum (and know that correlation coefficients extend from -1.00 to $+1.00$), you understand what a correlation matrix is, and you have considered several different kinds of bivariate correlation. Before you assume that you know everything there is to know about measuring the relationship between two variables, I'd like to provide you with six warnings. These warnings deal with the issue of cause, the coefficient of determination, the possibility of outliers, the assumption of linearity, the notion of independence, and criteria for claims of high and low correlations.

Correlation and Cause

It is important for you to know that a correlation coefficient does not speak to the issue of **cause and effect.** In other words, whether a particular variable has a causal impact on a different variable cannot be determined by measuring the two variables simultaneously and then correlating the two sets of data. Many recipients of research reports (and even a few researchers) make the mistake of thinking that a high correlation implies that one variable has a causal influence on the other variable. To prevent yourself from making this mistake, we suggest that you memorize this simple statement: correlation $\neq$ cause.

Consider Excerpt 3.28. In this excerpt, the researchers point out that their correlational findings should *not* be interpreted to mean that clear-cut cause-and-effect connections have been identified. Whereas these researchers set a good example of how researchers should talk about their correlational results, it is unfortunately the case that not all researchers follow their lead. Stated more forcefully, you need to be on guard for the many instances where researchers wrongfully impute "cause" into their correlational findings.

Later in this book, you will learn how researchers often collect data in such a way as to address the issue of cause. In such situations, however, researchers typically use data-gathering strategies that help them assess the possibility that one variable actually has a determining influence on a second variable. Those strategies

EXCERPT 3.28 • *Correlation and Cause*

Finally, it should be noted that correlation does not imply causation. The cross-sectional design of our study does not allow definitive causal inferences concerning the role of disturbed attachment patterns in the development of BPD.

Source: Fossati, A., Feeney, J. A., Carretta, I., Grazioli, F., Milesi, R., Leonardi, B., and Maffei, C. (2005). Modeling the relationship between adult attachment patterns and borderline personality disorder: The role of impulsivity and aggressiveness. *Journal of Social & Clinical Psychology, 24*(4), pp. 531–532.

require a consideration of issues that cannot be discussed here; in time, however, I am confident that you will come to understand the extra demands that are placed on researchers who want to investigate causal connections between variables. For now, all I can do is ask that you believe me when I say that correlational data alone cannot be used to establish a cause-and-effect situation.

Coefficient of Determination

Research Navigator.com
Coefficient of determination

To get a better feel for the strength of the relationship between two variables, many researchers will square the value of the correlation coefficient. For example, if r turns out equal to .80, the researcher will square .80 and obtain .64. When r is squared like this, the resulting value is called the **coefficient of determination.** In Excerpts 3.29 and 3.30, we see two research reports in which r^2 is presented. In the first of these excerpts, we see r^2 defined as the coefficient of determination. In Excerpt 3.30, we

EXCERPTS 3.29–3.30 • *The Coefficient of Determination*

Pearson product-moment correlation coefficient (r) was employed for correlational analysis. The coefficient of determination (r^2) was used to measure the meaningfulness of the relationship.

Source: Smitz, L. L., and Woods, A. B. (2005). Prevalence, severity, and correlates of depressive symptoms on admission to inpatient hospice. *Journal of Hospice & Palliative Nursing, 8*(2), p. 88.

[Results indicated that] participants who scored high on adherence were significantly more likely to have controlled blood pressure than were participants who scored low on adherence ($r = .58$). The coefficient of determination (r^2), indicating strength of the relationship between the Medication Adherence Scale and blood pressure control, was .33.

Source: Hill-Briggs, F., Gary, T. L., Bone, L. R., Hill, M. N., Levine, D. M., and Brancati, F. L. (2005). Medication adherence and diabetes control in urban African Americans with Type 2 diabetes. *Health Psychology, 24*(4), pp. 350, 351.

see that an r of .58 was "downsized" to .33 (by converting it into a coefficient of determination) in the sentence that discusses the "strength of the relationship."

The coefficient of determination indicates the proportion of variability in one variable that is associated with (or explained by) variability in the other variable. The value of r^2 will lie somewhere between 0 and +1.00, and researchers usually multiply by 100 so they can talk about the *percentage* of explained variability. In Excerpt 3.31, we see an example of where r^2 was converted into a percentage. As this excerpt indicates, researchers sometimes refer to this percentage as the amount of variance in one variable that's "accounted for" by the other variable, or they sometimes say that this percentage indicates the amount of "shared variance."

EXCERPT 3.31 • *r^2 and Explained Variation*

The Wheelchair User's Shoulder Pain Index (WUSPI) . . . measures how shoulder pain has interfered with different daily activities, such as transferring, wheeling, and self-care. . . . The Brief Pain inventory (BPI) (Short Form) was used to assess the subject's general experience of overall body pain, not isolated to the shoulder joint. . . . A modest correlation existed between WUSPI and BPI (r = 0.35) for all subjects, collectively. This translates to shoulder pain accounting for 12 percent of the variance of average whole body pain.

Source: Sawatzky, B. J., Slobogean, G. P., Reilly, C. W., Chambers, C. T., and Hol, A. T. (2005). Prevalence of shoulder pain in adult- versus childhood-onset wheelchair users: A pilot study. *Journal of Rehabilitation Research & Development, 42*, p. 4.

As suggested by the material in Excerpt 3.31, the value of r^2 indicates how much (proportionately speaking) variability in either variable is explained by the other variable. The implication of this is that the raw correlation coefficient (that is, the value of r when not squared) exaggerates how strong the relationship really is between two variables. Note that r must be stronger than .70 in order for there to be at least 50 percent explained variability. Or, consider the case where r = .50; here, only one-fourth of the variability is explained.

Outliers

My third warning concerns the effect of one or more data points that are located away from the bulk of the scores. Such data points are called **outliers,** and they can cause the size of a correlation coefficient to understate or exaggerate the strength of the relationship between two variables. Excerpt 3.32 very nicely illustrates this point.

In contrast to the good example provided in Excerpt 3.32, most researchers fail to check to see if one or more outliers serve to distort the statistical summary of the bivariate relationships they study. You won't see many scatter diagrams in journal articles, and thus you will not be able to examine the data yourself to see if

EXCERPT 3.32 • *Outliers*

This [calculated *r* of .433] suggests a positive relationship between the frequency of a product advertisement and the number of requests for that product. However, this may be a spurious relationship as two outliers, Barbie and Action Man products, appear to be responsible for the strength of this relationship. If these two products are excluded from the calculation, then the Pearson correlation changes completely [to $r = .065$, suggesting] no relationship between the two variables. . . .

Source: Pine, K. J., and Nash, A. (2002). Dear Santa: The effects of television advertising on young children. *International Journal of Behavioral Development, 26*(6), p. 536.

outliers were present. Almost always, you will be given just the correlation coefficient. Give the researcher some extra credit, however, whenever you see a statement to the effect that the correlation coefficient was computed after an examination of a scatter diagram revealed no outliers (or revealed an outlier that was removed prior to computing the correlation coefficient).

Linearity

The most popular technique for assessing the strength of a bivariate relationship is Pearson's product-moment correlation. This correlational procedure works nicely if the two variables have a linear relationship. Pearson's technique does not work well, however, if a curvilinear relationship exists between the two variables.

A **linear** relationship does *not* require that all data points (in a scatter diagram) lie on a straight line. Instead, what *is* required is that the *path* of the data points be straight. The path itself can be very narrow, with most data points falling near an imaginary straight line, or the path can be very wide—so long as the path is straight. (Regardless of how narrow or wide the path is, the path to which we refer can be tilted at any angle.)

If a **curvilinear** relationship exists between two variables, Pearson's correlation will underestimate the strength of the relationship that is present in the data. Accordingly, you can place more confidence in any correlation coefficient you see when the researcher who presents it indicates that a scatter diagram was inspected to see whether the relationship was linear before Pearson's *r* was used to summarize the nature and strength of the relationship. Conversely, add a few grains of salt to the *r*s that are thrown your way without statements concerning the linearity of the data.

In Excerpts 3.33 and 3.34, we see two examples where researchers checked to see if their bivariate data sets were linear. The first of these excerpts illustrates how scatter diagrams can be used to accomplish this task. In Excerpt 3.34, we see that an inspection of a scatter diagram can provide a "red flag" that a relationship is not linear. The researchers associated with these two excerpts deserve high praise for taking the time to check out the linearity assumption before computing Pearson's *r*. Unfortunately, however, most researchers collect their data and compute correlation coefficients without ever thinking about linearity.

EXCERPTS 3.33–3.34 • *Linearity and Curvilinearity*

Pearson's product-moment correlations were used to examine the nature of the relationships between the subscales . . . and linearity between variables was assessed by inspection of bivariate scatter plots.

Source: Jurkovic, D., and Walker, G. A. (2006). Examining masculine gender-role conflict and stress in relation to religious orientation and spiritual well-being in Australian men. *Journal of Men's Studies, 14*(1), pp. 34, 35.

Inspection of scatter plots of outcome and treatment duration showed a curvilinear relationship.

Source: Lorentzen, S., and Høglend, P. (2005). Predictors of change after long-term analytic group psychotherapy. *Journal of Clinical Psychology, 61*(12), p. 1547.

Correlation and Independence

In many empirical studies, the researcher will either build or use different tests in an effort to assess different skills, traits, or characteristics of the people, animals, or objects from whom measurements are taken. Obviously, time and money will be wasted if two or more of these tests are redundant. Stated differently, it is desirable (in many studies) for each measuring instrument to accomplish something unique compared to the other measuring instruments being used. Two instruments that do this are said to be **independent.**

The technique of correlation often helps researchers assess the extent to which their measuring instruments are independent. Independence exists to the extent that r turns out to be close to zero. In other words, low correlations imply independence whereas high positive or negative correlations signal lack of independence.

Excerpt 3.35 illustrates how authors will sometimes use the term *independent* in their research reports. Note that it is *low* values of r that signal independence. Also note that it is the variables (and not the correlations themselves) that are considered to be independent; a low r is simply the "signpost" that suggests the presence of independence.

EXCERPT 3.35 • *Independence*

Parent and adolescent reports of parental monitoring correlated only .26 [thereby] indicating that the two measures are largely independent ($r^2 = .07$).

Source: Donenberg, G. R., Wilson, H. W., Emerson, E., and Bryant, F. B. (2002). Holding the line with a watchful eye: The impact of perceived parental permissiveness and parental monitoring on risky sexual behavior among adolescents in psychiatric care. *AIDS Education and Prevention, 14*(20), p. 153.

Relationship Strength

My final warning concerns the labels that researchers attach to their correlation coefficients. There are no hard and fast rules that dictate when labels such as "strong" or "moderate" or "weak" should be used. In other words, there is subjectivity involved in deciding whether a given r is "high" or "low." Not surprisingly, researchers are sometimes biased (by how they *wanted* their results to turn out) when they select an adjective to describe their obtained rs. Being aware that this happens, you need to realize that you have the full right to look at a researcher's r and label it however you wish, even if your label is different from the researcher's.

Consider Excerpt 3.36. In this passage, the researchers assert that two of the obtained rs indicate "a strong correlation." Knowing now about how the coefficient of determination is computed and interpreted, you ought to be a bit hesitant to swallow the researchers' assertion that .396 and .310 are strong correlations. If squared and then turned into percentages, these rs indicate that (1) less than 16 percent of the variability in TV watching time was associated with BMI (body mass index) and (2) less than 10 percent was associated with triceps skinfold thickness.

EXCERPT 3.36 • *Questionable Labels Used to Describe Relationships*

Pearson's correlation analysis was used to correlate television watching and body composition. . . . When the correlation was investigated between TV watching time and weight, BMI, skinfold thickness, waist to hip ratio and %BF, it was seen that there was a strong correlation between TV watching time and BMI [$r = 0.396$] and triceps skinfolds [$r = 0.310$].

Source: Özdirenç, M., Özcan, A., Akin, F., and Gelecek, N. (2005). Physical fitness in rural children compared with urban children in Turkey. *Pediatrics International, 47*(1), pp. 28, 29.

Review Terms

Abscissa
Biserial correlation
Bivariate
Cause and effect
Coefficient of determination
Correlation coefficient
Correlation matrix
Cramer's V
Curvilinear
Dichotomous variable
Direct relationship
High
Independent
Indirect relationship
Inverse relationship
Kendall's tau
Linear
Low
Moderate
Negative correlation
Nominal
Ordinal

Ordinate
Outlier
Pearson's product-moment correlation
Perfect
Phi
Point biserial correlation
Positive correlation
Qualitative variable
Quantitative variable
Rank-order correlation
Ranks
Raw score
Relationship
Scatter diagram
Spearman's rho
Strong
Tetrachoric correlation
Weak
r
r_s
r^2
r_{pb}
r_{bis}
V
ρ

The Best Items in the Companion Website

1. An interactive online quiz (with immediate feedback provided) covering Chapter 3.
2. Ten misconceptions about the content of Chapter 3.
3. The author's poem "True Experimental Design."
4. An email message from the author to his students in which he asks an honest question about Pearson's *r* and Spearman's rho.
5. Two jokes about statistics, the first of which concerns a student in a statistics course.

To access chapter objectives, practice tests, weblinks, and flashcards, visit the companion website at www.ablongman.com/huck5e.

Fun Exercises inside Research Navigator

Research Navigator.com

1. Is IQ correlated with reaction time?

The 81 adult participants in this study were measured on two kinds of reaction time: how fast they could respond to auditory stimuli and how fast they could respond to visual stimuli. They were also measured on several other variables, one of which was IQ. The correlation between the two kinds of reaction time turned out as most people would expect—it was moderate and positive, with Pearson's $r = .69$. But what about the correlation between IQ and each of the two kinds of reaction time? These two *r*s were about the same size and they had the same sign. But what was that sign? In other words, do you think this study revealed a direct relationship or an indirect relationship between IQ and reaction time? After you make your guess, refer to the PDF version of the research report in the Communication database of Content-Select and look at the correlation matrix at the bottom of page 46.

R. Stringer & K. E. Stanovich. The connection between reaction time and variation in reading ability: Unraveling covariance relationships with cognitive ability and phonological sensitivity. *Scientific Studies of Reading*. Located in the COMMUNICATION database of ContentSelect.

2. How does emotional intelligence correlate with general intelligence?

In this study, 107 undergraduate students, graduate students, and college graduates were measured in terms of their general intelligence. In addition, the research participants completed the *Multifactorial Emotional Intelligence Scale*. This instrument provided four scores per participant, three based on its subscales (Emotional Knowledge, Emotional Perception, and Emotional Regulation) and one composite score. The correlation matrix included in the research report contained 21 bivariate correlation coefficients, four of which showed the relationship between the study's measures of general intelligence and the four scores on emotional intelligence. These four *r*s turned out equal to .40, .23, .04, and −.03. Which emotional intelligence score do you think correlated most highly with general intelligence? After making a guess, locate the PDF version of the research report in the Psychology database of Content-Select and look at Table 1 at the bottom of page 188.

J. Pellitteri. The relationship between emotional intelligence and ego defense mechanisms. *Journal of Psychology*. Located in the PSYCHOLOGY database of ContentSelect.

Review Questions and Answers begin on page 513.