

of measures we have just been considering—percentiles and standard scores—provide a framework with which we can view the performance of a specific person. Both provide a way to view the person's performance relative to a specific *reference group*. Percentile information tells us what point in the score distribution just exceeds a specified fraction of the scores for the group; the standard deviation provides a common unit of distance from the mean of the group. The individual's score can then be expressed as a distance above or below the mean in these common units. In this way it is possible to give a meaningful answer to questions such as "Are you taller than you are heavy?"

MEASURES OF RELATIONSHIP

We look now for a statistic to express the relationship between two sets of scores. For example, in Table 2-1, for each pupil we have scores for reading, mathematics, and spelling. To what extent did those pupils who scored well in mathematics also score well on the reading test? In this case, we have two scores for each individual. We can picture these scores using a graph in two dimensions, one dimension for each test. Such a graph is shown in Figure 2-9 and is called a **scatterplot**. The first person listed in Table 2-1, Aaron Andrews, had a reading test score of 32 and a mathematics test score of 43. His scores are represented by the * in Figure 2-9, plotted at 32 on the vertical, or reading, scale and at 43 on the horizontal, or mathematics, scale. The dots in the figure each represent one of the other 51 pupils' paired scores. For example, the dot at a reading score of 21 and a math score of 19 represents Quadra Quickly's scores.

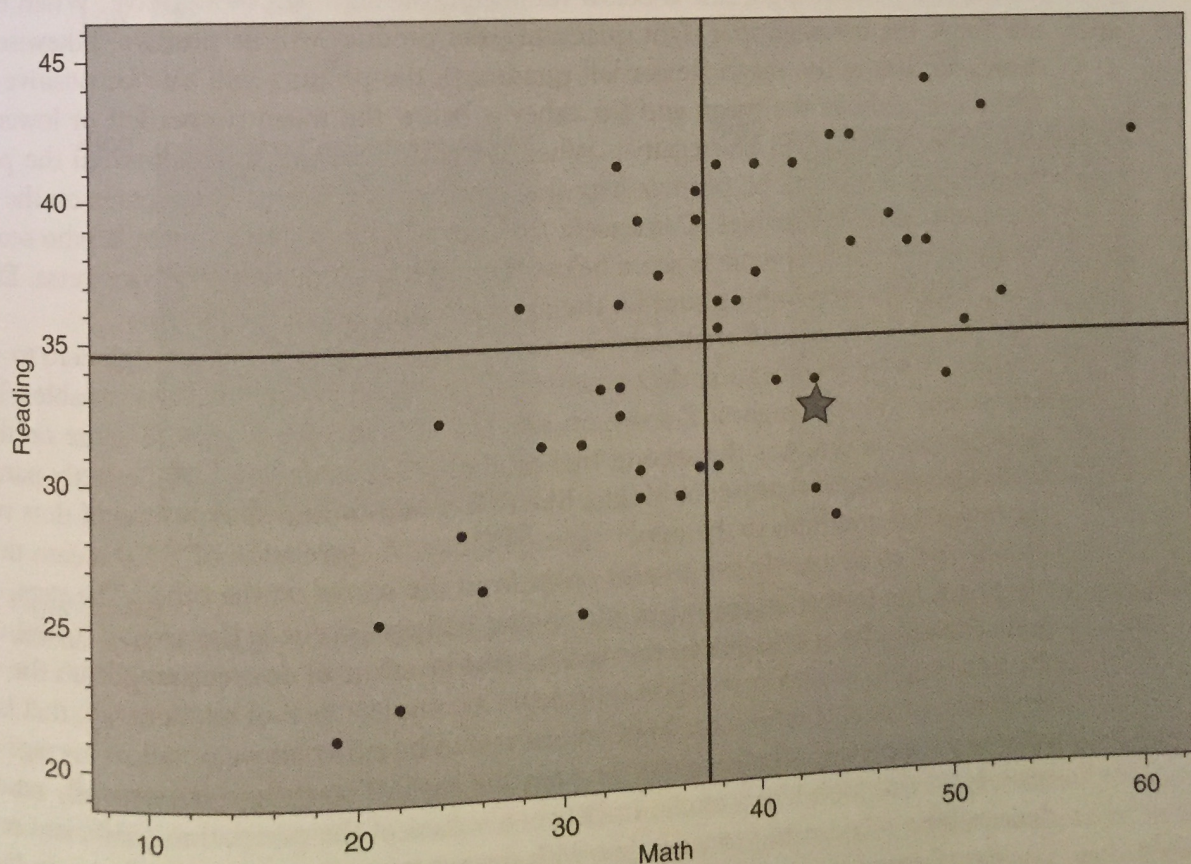


Figure 2-9
Scatterplot of reading and mathematics scores.

The vertical and horizontal lines drawn through the scatterplot are the means of the two variables and divide the plot into four parts, or **quadrants**. When a person who does well in reading (above the mean) also does well in mathematics, the dot representing that pair of scores falls in the upper right-hand quadrant. The dot for one who does poorly (below the mean) on both tests falls in the lower left quadrant. Where a good score on one test is paired with a poor score on the other, we find the points falling in the other quadrants, that is, the upper left and the lower right. Students who score in the middle on both tests are represented by points in the center of the plot. Inspection of Figure 2-9 reveals some tendency for the scores to scatter in the lower left to the upper right direction, from low reading *and* low mathematics to high reading *and* high mathematics, but there are many exceptions. The relationship is far from perfect; it is a matter of degree. We need some type of index to express this degree of relationship.

As an index of degree of relationship, a statistic known as the **correlation coefficient** is widely used. The symbol r is used to designate this coefficient. Looking at the formula for r will help us understand why the coefficient has the properties it does. The correlation coefficient is defined as

$$r = \frac{\sum Z_X Z_Y}{N}$$

where Z_X and Z_Y are the pair of standard scores for an individual (such as the +0.54 and -0.44 scores for Aaron Andrews). What the formula tells us to do is multiply each person's standard score on one variable by their standard score on the other, sum the products across all cases, and divide the result by the number of people in our sample.

Now think about those pairs of standard scores. If a person's Z-score is above the mean, it will have a positive sign; if it is below the mean, the sign will be negative. When both Z-scores are above the mean (upper right quadrant), the product will be positive. Likewise, when both scores are below the mean (lower left quadrant), the product will also be positive. However, if one score is above the mean and the other is below the mean (upper left or lower right quadrant), the product will be negative. When we add these products across all the people in our sample, the result will be positive if people tend, on average, to have scores on the same side of the mean on both variables. Conversely, the sum will be negative if people who score above the mean on one variable tend to score below the mean on the other, and vice versa. Dividing by N simply adjusts the resulting sum for the number of people in the group.

The correlation coefficient can take values ranging from +1.0 through zero to -1.0. A correlation of +1.0 signifies a perfect positive relationship between the two variables. It means that the person with the highest Z-score on one test also had the highest Z-score on the other, the next highest on one was the second highest on the other, and so forth, exactly parallel through the whole group. A scatterplot of data like this would form a straight line of dots running from the lower left quadrant to the upper right quadrant. A correlation of -1.0 means that the scores on one test go in exactly the reverse order from the scores on the other. The person highest on one test is the lowest on the other, the second highest on one is the second lowest on the other, and so forth. The scatterplot in this case would be a line of dots running from the upper left to the lower right. A zero correlation represents a complete lack of relationship; that is, there is no tendency for people who score high on one test to be either above or below average on the other (the positive products balance out the negative ones). The pattern is essentially random and the scatterplot will look like a circle. In-between values of the correlation coefficient represent tendencies for a relationship to exist, but with discrepancies, as in the case shown in Figure 2-9.

Every correlation coefficient contains two pieces of information. One is the sign of the correlation, which tells whether the two variables tend to rank people in the same order (plus) or in

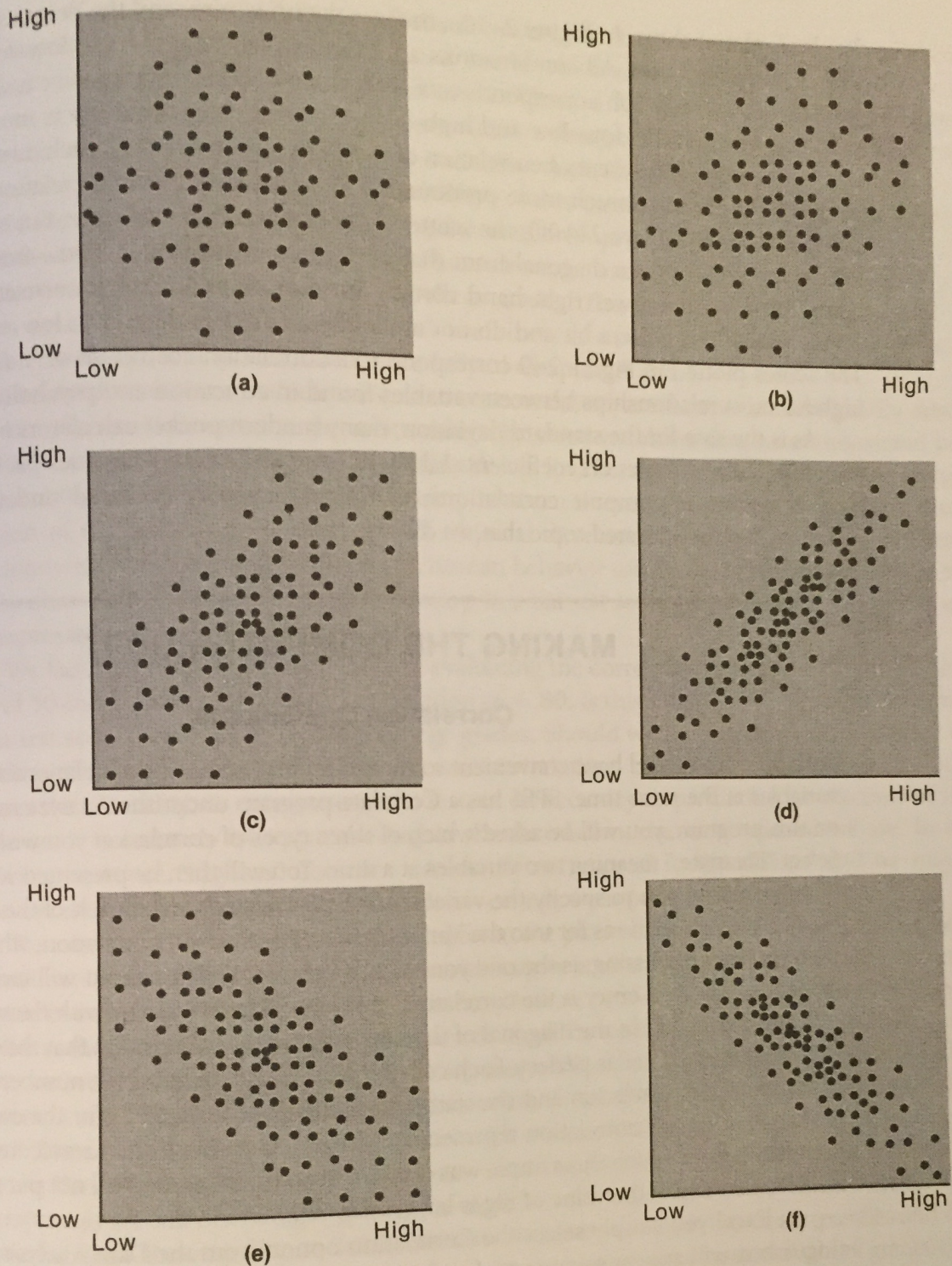


Figure 2-10 Distribution of scores for representative values of correlation coefficient: (a) correlation of 0.00, (b) correlation of +.30, (c) correlation of +.60, (d) correlation of +.90, (e) correlation of $-.60$, and (f) correlation of $-.90$.

the reverse order (minus). The second piece of information is the *magnitude* of the correlation, which tells how strong the relationship is. The correlations $+.50$ and $-.50$ indicate the same strength of relationship, but the first reveals that there is some tendency for people to be in the same rank order on both variables; the second shows some tendency for people with higher scores on one variable to have lower scores on the other. Figure 2-10 illustrates four different

levels of relationships. In Figure 2-10a, the correlation is zero, and the points scatter in a pattern that is almost circular. All combinations are found equally: high-high, low-low, high-low, and low-high. Figure 2-10b corresponds to a correlation of $+ .30$. You can see a slight trend for the points to group in the low-low and high-high direction. This tendency is more marked in Figure 2-10c, which represents a correlation of $+ .60$. In Figure 2-10d, which portrays a correlation of $+ .90$, the trend is much more pronounced. Note that when the correlation is $- .60$ (Figure 2-10e) or $- .90$ (Figure 2-10f), the scattering of the points is the same, but the swarm of dots falls along the opposite diagonal from that of Figures 2-10c and 2-10d—from the upper left-hand corner to the lower right-hand corner. But even with as high a correlation as $+ .90$, the scores spread out quite a bit and do not all lie directly on the line from low-low to high-high. The scores plotted in Figure 2-9 correspond to a correlation coefficient of $+ .62$, which is fairly high for most relationships between variables found in education and psychology.

As is the case for the standard deviation, many modern pocket calculators include a program to compute the correlation coefficient. Likewise, spreadsheet and statistics packages will always have programs to compute correlations, although they may be listed under the heading of *regression*, a closely related topic that we discuss next.

MAKING THE COMPUTER DO IT

Correlation Coefficients

Both SPSS and Excel have convenient routines for computing the correlations among several variables at the same time. SPSS has a Correlate program under the Analyze menu. Clicking on this program, you will be asked which of three types of correlation you wish to compute. Select "Bivariate," meaning two variables at a time. You will then be presented with a dialogue box that allows you to specify the variables to be correlated. Move each of the variables you want to find correlations for into the Variables box. The Pearson correlation, which is the one we have been discussing, is the one you want, so click OK. The output will include a square table in which each entry is the correlation of the variable in the row with the variable in the column. The values in the diagonal of the table will be 1.0, which says that the correlation of any variable with itself is perfect. Each cell of the table also includes the number of cases used to compute the correlation and the statistical significance ("Sig") test for the correlation (the probability that this correlation represents a random deviation from a zero correlation in the population from which this sample was drawn). Most investigators will not put much faith in a correlation where the value of Sig is larger than .05.

In Excel you simply select the Correlation option from the Data Analysis menu. A dialogue box will appear asking you for the Input Range. You must highlight all of the scores to be included in the analysis, or specify the upper left cell and the lower right cell of the part of the table containing the scores you want to analyze. Then click on the "Output Range" button in the dialogue box, click in the "Output Range" box, and click in an empty cell where you wish the upper left corner of the correlation table to go. The output will be the lower half of the table you would get from SPSS, including the ones in the diagonal. Excel also has a Correl function on the Σ or f_X menu that you can use to find the correlation between a pair of variables. Highlight the scores on one variable for Array 1, then highlight the scores on the other variable for Array 2. Click OK, and the correlation will appear in the cell into which you have pasted the function.

You will encounter correlation coefficients in connection with testing and measurement in three important settings. The first situation is one in which we are trying to determine how precise and consistent a measurement procedure is. Thus, if we want to know how consistent a measure of speed we can expect to get from runners doing a 50-meter dash, we can have each person run the distance twice, perhaps on successive days. Correlating the two sets of scores will give information on the stability, or **reliability**, of this measure of running speed. The second situation is one in which we are studying the relationship between two different measures to evaluate one as a **predictor** of the other. Thus, we might want to study a scholastic achievement test from high school as a predictor of college grades. The correlation of the test scores with grades would give an indication of the test's usefulness as a predictor. These two uses of the correlation coefficient will be described more fully in Chapters 4 and 5.

The third situation in which we encounter correlation coefficients is more purely descriptive. We often are interested in the relationships between variables, simply to understand better how behavior is organized. What correlations do we find between measures of verbal and quantitative abilities? How close is the relationship between interest in mechanical jobs and comprehension of mechanical devices? Is rate of physical development related to rate of intellectual development? Many research problems in human behavior can best—or perhaps only—be studied by observing relationships as they develop in a natural setting, and these relationships often are expressed with correlation coefficients.

We face the problem, in each case, of evaluating the correlation we obtain. Suppose the two sets of 50-meter dash scores yield a correlation of $+ .80$. Is this satisfactory? Suppose the achievement test scores correlate $+ .60$ with college grades. Should we be pleased or discouraged? (The wording of this question implies that we want correlations to be high. When a correlation expresses the reliability, or consistency, of a test, or its accuracy in predicting an outcome of interest to us, it is certainly true that the higher the correlation, the more pleased we are. In other contexts, however, "bigness" does not necessarily correspond to "goodness," and we may not have a preference on the size of a correlation—or may even prefer a low one.)

The answer to the third question lies in part in the plots of Figure 2-10. Clearly, the higher the correlation, the more closely scores on one variable agree with scores on the other. If we think of discrepancies away from the diagonal line, from low-low to high-high, as "errors," the errors become smaller as the correlation becomes larger. If we think of the *standard deviation of scores on test Y* as the errors we would make in predicting people's scores without test X, then the correlation coefficient between X and Y tells us how much our errors will be reduced by using test X as a predictor of variable Y. The square of the correlation coefficient (r^2) tells us how much our errors in estimating people's test Y scores are reduced using test X over what those errors would be without the test. But these discrepancies are still discouragingly large for even rather substantial correlation coefficients, for example, for those shown in Figure 2-10. We must always be aware of these discrepancies and realize that even with a correlation such as $+ .60$ between achievement test scores and school grades (which is about as high as correlations between these measures usually get), there still will be a number of students whose school performance will differ a good deal from the best prediction we can make from the test. We will discuss these issues in more detail in Chapter 5.

However, everything is relative, and any given correlation coefficient must be interpreted in comparison with values that are commonly obtained. Table 2-6 contains a number of correlations that have been reported for different types of variables. The nature of the scores being correlated is described, and the coefficient is reported. An examination of this table will provide some initial background for interpreting correlation coefficients. The correlation coefficient will gradually take on added meaning as you encounter coefficients of different sizes in your reading and as you work with tests.

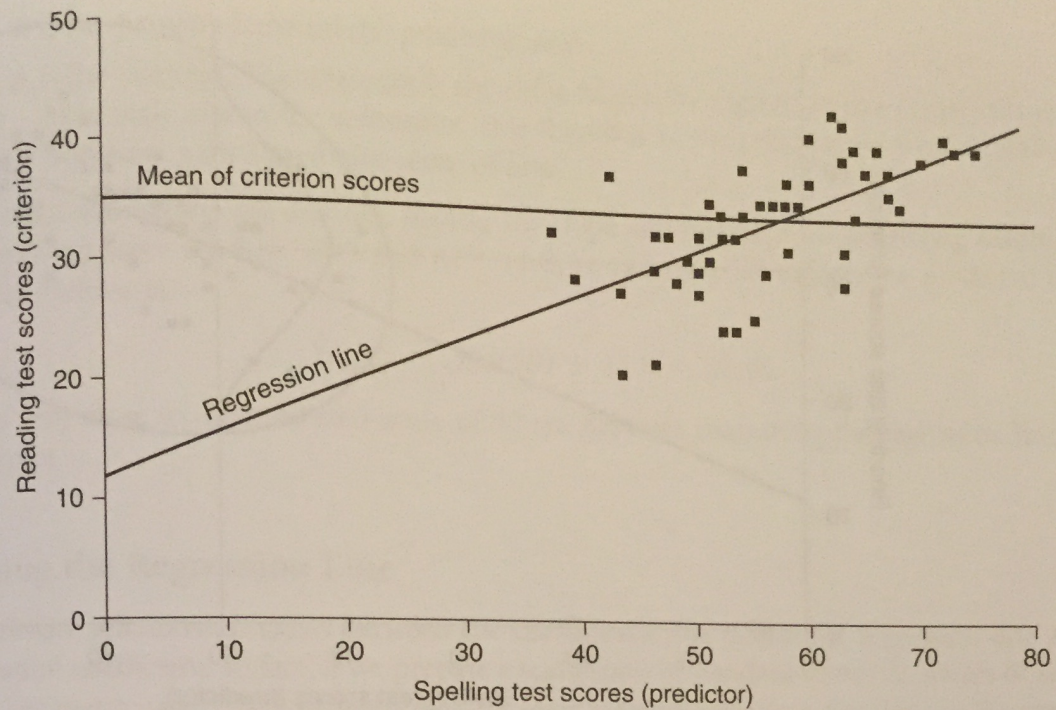
Table 2-6
Correlations Between Selected Variables

Variable	Correlation Coefficient
Heights of identical twins	.95
Intelligence test scores of identical twins	.88
Reading test scores in Grade 3 and Grade 6	.80
Rank in high school class and teacher's rating of work habits	.73
Height and weight in 10-year-olds	.60
Arithmetic computation test score and nonverbal intelligence test (Grade 8)	.54
Height of brothers (adjusted for age)	.50
Intelligence test score and parents' occupational level	.30
Strength of grip and running speed	.16
Adult height and intelligence test score	.06
Ratio of head length to width and intelligence test score	.01
Armed Forces Qualification Test scores of recruits and number of school grades repeated	-.27
Artist interest score and banker interest score	-.64

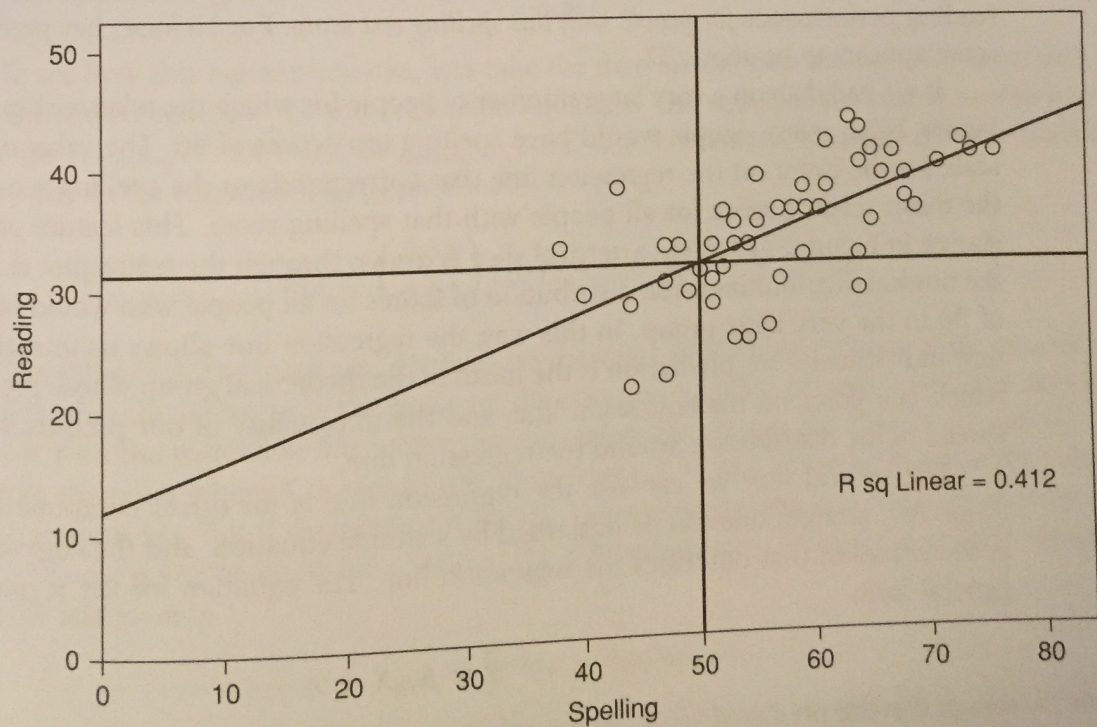
MAKING PREDICTIONS

We have seen that the correlation coefficient tells us the degree of association between two variables and the direction of that relationship. We can use this information, along with the means and standard deviations of the predictor and criterion variables, to answer our seventh question, how can we make the best possible prediction about criterion performance for each person, based on his or her score on the predictor? To do this we use what is called a **regression equation**. The regression equation tells us what our best guess of a person's score on an outcome variable, known as the criterion, would be, given the person's score on the predictor.

To see how the regression equation works, let us look at the scatterplot of the relationship between scores on the reading test and scores on the spelling test for Mr. Cordero's and Ms. Johnson's sixth-grade classes. Suppose we wish to use the spelling test to predict what level of reading proficiency we might expect from Heloise Abelard, a student who is being considered for transfer into this school. The scatterplot for the combined classes is shown in Figure 2-11. We have drawn two lines through the scatterplot. One is a horizontal line that shows the mean reading test score for all students for whom we have test scores. This is our best guess about Heloise's likely reading level if we *know nothing else about her*. In general, the mean of the criterion scores will always be our best guess about a person's score if we don't have any additional information. The regression equation is designed to allow us to make optimal use of any additional information we may have. What the regression equation does is describe the other line through the scatter plot, called the **regression line**. This line results in the most accurate predictions we can make *using the information from our predictor*. We can use the regression line in either of two ways. The first is visual and is illustrated in Figure 2-12. Suppose Heloise earned a score of 50 on the spelling test. If we draw a line from the

**Figure 2-11**

Scatterplot of scores from the reading and spelling tests given by Mr. Cordero and Ms. Johnson showing the overall mean of criterion scores and the regression line.

**Figure 2-12**

Using the regression line visually to make the best prediction of reading test score for Heloise.

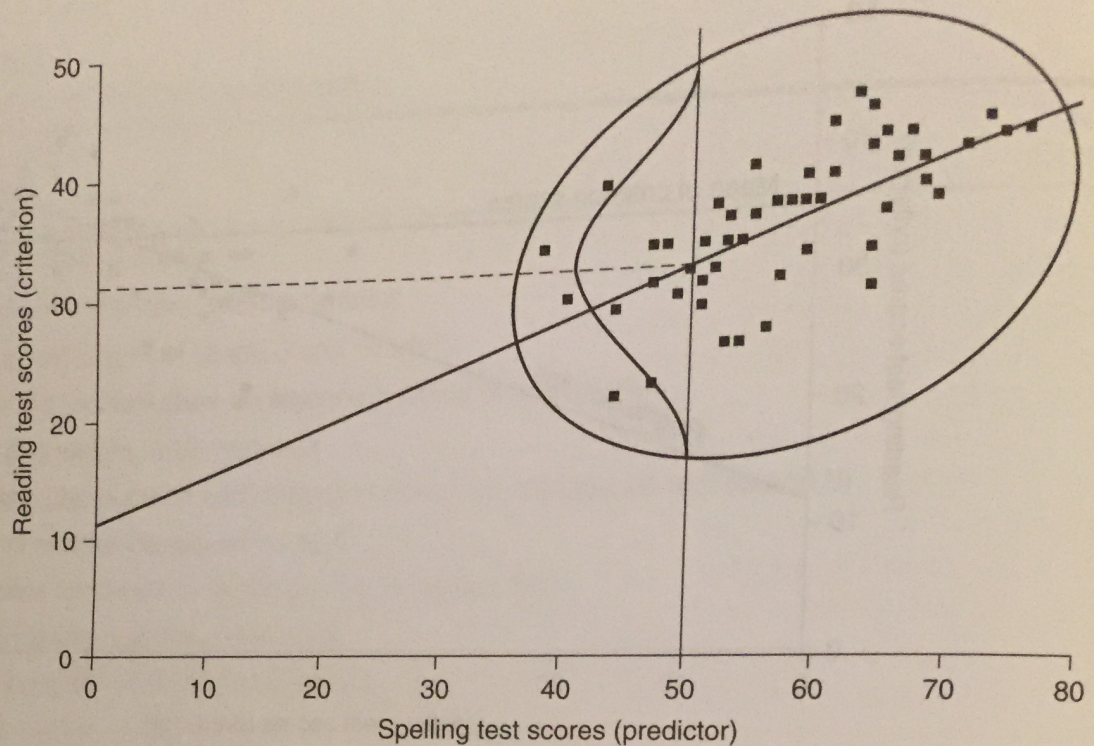


Figure 2-13

Scatterplot of reading and spelling test scores for a large group showing the distribution of reading test scores for all people who earned spelling scores of 50. The standard deviation of this distribution is the standard error of estimate.

spelling test score of 50 directly up to the regression line, and then horizontally over to the scale for the reading test, the value on the reading test scale will be the best prediction we can get of reading performance *for people with this spelling test score*. For Heloise, her predicted reading test score appears to be about 32.

If we had data on a very large number of people for which the relationship resembled that in Figure 2-12, many people would have spelling test scores of 50. The value on the reading test scale for the point on the regression line that corresponds to the spelling score of 50 would be the mean reading score for all people with that spelling score. This feature of the scatterplot is shown in Figure 2-13. Here a vertical slice is drawn through the scatterplot at a score of 50, and the normal distribution is the distribution of scores for all people who would have earned a score of 50 in the very large group. In this way, the regression line allows us to make predictions for new individuals. Our prediction is the mean of the theoretical group whose predictor score is 50, which is a point on the regression line, and the uncertainty of our prediction is shown by the spread of the distribution around the regression line.

The second way we can use the regression line is for direct computation of a predicted score. Any straight line can be described by a simple equation, and the **regression equation** is the equation that describes the regression line. The equation for the regression line has the general form

$$\hat{Y} = B_{YX}X + A$$

where $\hat{Y}$ is the predicted score on the criterion,

B_{YX} is the slope of the regression line for predicting Y from X (the slope is the number of units that the line rises or falls in the scale of the criterion variable for every unit of increase we have in the predictor),

X is the person's score on the predictor, and

A is the *intercept*. The intercept is the value where the regression line crosses the criterion scale axis in the scatterplot. It is the value of the criterion we would predict for someone with a predictor score of zero.

Using relationships we will describe shortly, the slope and intercept for predicting reading score from spelling score are $B = .393$ and $A = 11.8$, respectively. Therefore, the predicted reading score for Heloise is

$$\hat{Y}_{\text{Heloise}} = .393(50) + 11.8 = 31.45$$

This is very close to the predicted score of 32 we got from examining the regression line in the scatterplot.

Finding the Regression Line

An intimate relationship exists between the coefficients that define the regression line and the correlation coefficient. In fact, if we prepare a scatterplot of standard scores (Z-scores or scores in any other metric where *both variables have the same standard deviation*), the slope of the regression line is the correlation coefficient. That is, when $SD_Y = SD_X$,

$$B_{YX} = r_{YX}$$

However, when the two standard deviations are not equal, B is related to r by the ratio of the SDs. Specifically, when we are predicting Y from X , the relationship is

$$B_{YX} = \left(\frac{SD_Y}{SD_X} \right) r_{YX}$$

To see how this equation works, let's take the data for our two sixth-grade classes and compute the slope coefficient for our regression equation. The standard deviations for the two variables are $SD_{\text{Spelling}} = 9.04$ and $SD_{\text{Reading}} = 5.55$. The correlation between the two variables is $r_{YX} = .64$, so the equation for B_{YX} produces

$$B_{YX} = \left(\frac{5.55}{9.04} \right) (.64) = (.61)(.64) = .393$$

The value of A , the intercept, depends on the value of B and the means of the two variables. Recalling that the intercept is the point on the criterion variable (Y) axis when X is zero (or the value of Y on the regression line where the regression line crosses the Y axis), we can see that to get from the point where X is zero to the point where $X = M_X$, we must increase the value of X by the amount M_X . But if we increase X by M_X , the regression line is going to change by $B(M_X)$. Since the only point we know for certain in the Y distribution is the mean of Y (M_Y), we can find A by the relationship

$$A = M_Y - B_{YX}M_X$$

In words, we can say that the intercept is the mean of the criterion variable minus (B times the mean of the predictor variable). For the two variables in our example we have $M_{\text{Reading}} = 34.44$ and $M_{\text{Spelling}} = 57.54$. The equation for A therefore produces a value of

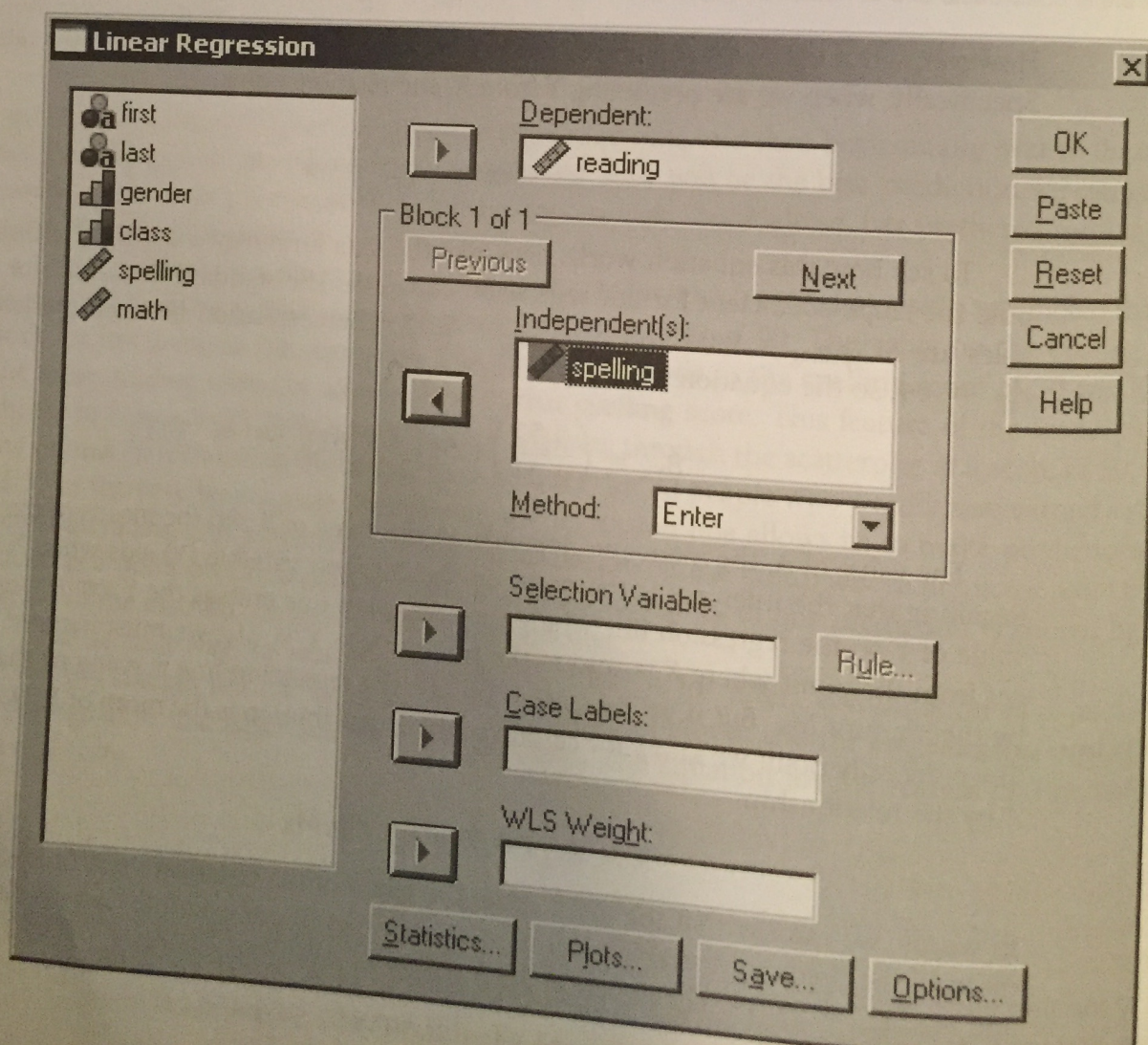
$$A = 34.44 - [(.393)(57.54)] = 11.8$$

which agrees with the graph in Figure 2-12. The equation $\hat{Y} = .393X + 11.8$ can be used to make predictions for as many new students as we wish. The same equation would be used for all new students until we have collected criterion data on them, at which point we might add them to our sample and recompute the values for the regression equation.

MAKING THE COMPUTER DO IT

Regression

It is quite simple to get the regression equation from either SPSS or Excel. Both have a program called Regression that will compute the necessary values. In SPSS the Regression program is under the Analysis menu. The program offers several types of regression, but the one we want is called "Linear." When you see the following screen, simply place the criterion variable in the Dependent box and the predictor in the Independent box. You can use the Statistics button to request additional output such as the means and standard deviations (under Descriptives). Click OK and you will get the regression slope and intercept. The value for the intercept is in the row labeled **Constant**, and the value for the predictor will be in the row with its label. The program can handle many predictors simultaneously, but how this is done is beyond our scope.



Regression

Input

Input Y Range:

Input X Range:

☐ Labels

☐ Constant is Zero

☐ Confidence Level: %

Output options

☒ Output Range:

☐ New Worksheet Ply:

☐ New Workbook

Residuals

☐ Residuals

☐ Standardized Residuals

☐ Residual Plots

☐ Line Fit Plots

Normal Probability

☐ Normal Probability Plots

OK Cancel Help

Excel has a regression program under the Data Analysis option on the Tools menu. When you select this option, you will be prompted to "Input Y range" and "Input X range." Y is the criterion and you specify the cells in which the Y scores are located. Do the same for X. If you wish the results to be displayed in the same spreadsheet as the data, select Output Range and specify a cell in which the output should start, usually a cell to the right of the data. For our example, the screen should look like the screen above.

If you use either program to replicate the analysis we have described in this section, you will get slightly different values than we did when these were calculated by hand. Both programs output values of $B = .394$ and $A = 11.76$. The reason for the difference is that the programs keep more decimal places in their intermediate calculations. It will often be the case that hand calculations will differ slightly from the computer for this reason.

SUMMARY

All measurements produce numbers. It is usually desirable to summarize the numbers that result from measurements so that we can both answer specific questions about the people we have measured and draw general conclusions. We have focused on seven basic types of questions.

1. What type of scale does our measurement represent?

Different uses of numbers allow us to do different things with the resulting data. In nominal scales, numbers substitute for labels. Ordinal scales are usually in the form of ranks. Interval

and ratio scales allow us to find means and standard deviations.

2. *What does the general array of numbers "look like"? Or how do they arrange themselves?*
The answer to this question involves sorting the scores into a frequency distribution (Table 2-2) or a grouped frequency distribution (Table 2-3). Histograms (Figure 2-1) can also provide a picture of the data.
3. *What does the typical individual look like? Or where is the middle of the group?*
The mode, median (or 50th percentile), and mean are indices of where the middle of the group falls. The mean, or arithmetic average, is the most commonly used measure of the center of the group.
4. *How widely do the scores spread out around their center? Or what is their spread?*
The spread of scores can be represented by the range, the semi-interquartile range (half the distance between the 25th and the 75th percentiles), and the standard deviation, which is a measure of how far scores deviate, on the average, from the mean.
5. *How should we interpret the score of any individual?*
We will discuss this topic in Chapter 3.

6. *To what extent do the scores from two measurements "go together"? Or what is the degree of relationship between the two sets of scores?*
Generally, the most useful measurement of relationships is the correlation coefficient. This index runs from +1.0, indicating a perfect positive relationship or exact agreement, through zero, indicating no association between the sets of scores, to -1.00, indicating a perfect negative relationship, or exact disagreement, on the order of individuals on the two traits. The correlation coefficient is important as an index of stability of measurement and as a measure of how well one trait can be predicted from another.
7. *How can we make the most accurate possible prediction of a person's performance on one measurement from their performance on another measurement?*
The regression line provides the most accurate prediction of a person's score on variable Y from their score on variable X. The regression line is defined by its slope, B, or the number of units that Y changes for every unit of change in X, and its intercept, A, or the value of Y when X is zero. The slope of the regression line has the same sign as the correlation coefficient, but is affected by the relative sizes of the two variables' standard deviations.

QUESTIONS AND EXERCISES

1. For each of the following sets of scores, select the most suitable score interval and set up a form for tallying the scores:

Test	Number of Cases	Range of Scores
Mathematics	103	15-65
Reading Comprehension	60	60-140
Interest Inventory	582	65-248

intervals (i.e., the dividing points between them):

a.	b.	c.
4-7	17-19	60-69
8-11	20-22	70-79
12-15	23-25	80-89
.	.	.
.	.	.

2. In each of the following distributions, indicate the size of the score interval, the midpoints of the intervals shown, and the real limits of the

3. Using the spelling scores given in Table 2-1, make a frequency distribution and a histogram. Compute the median and the upper and lower

quartiles. Also, compute the arithmetic mean and the standard deviation from the original scores.

4. The Bureau of Census uses the median in reporting average income. Why is this index used rather than the mean?
5. A 50-item mathematics test was given to the 150 students in five classes at Sunnyside School. Scores ranged from 16 to 50, with 93 of the students getting scores above 40. What would this score distribution look like? What could you say about the suitability of this test for this group? What measures of central tendency and variability would be most suitable? Why?
6. A high school teacher gave two sections of a biology class the same test. The results were as follows:

	Section A	Section B
Median	74.6	74.3
Mean	75.0	73.2
75th percentile	79.0	80.0
25th percentile	71.0	64.4
Standard deviation	6.0	10.5

From these data, what can you say about these two classes? What implications do the test results have for teaching the two groups?

7. A test in history given to 2500 10th-grade students had a mean of 52 and a standard deviation of 10.5. How many standard deviations above or below the mean would the following students fall?

Heather	48	Rob	60	Marc	31
Krista	56	Tina	36	Bill	84

8. If the distribution in Question 7 was approximately normal, what percentage of the group would Heather, Krista, Rob, Tina, Marc, and Bill each surpass?
9. Assuming that a set of scores is normally distributed with a mean of 82 and a standard deviation of 12, what would be the percentile rank for each of the following scores?
 - a. 74
 - b. 85
 - c. 99
10. Explain the meaning of each of the following correlation coefficients:
 - a. The correlation between scores on a reading test and scores on a group test of general intellectual ability is $+0.78$.
 - b. Ratings of students on good citizenship and on aggressiveness show a correlation of -0.56 .
 - c. The correlation between weight and sociability is $+0.02$.
11. Jason's score on the Thorndike Precarious Prognosticator of Potential Proficiency (TP⁴) is 87. The test has a mean of 100 and a standard deviation of 20. The mean freshman grade point average at Flunkemout University is 2.5 with a standard deviation of .5. The correlation between the TP⁴ and grades at FU is $-.60$.
 - a. What are the slope and intercept of the regression equation (be careful; watch the sign)?
 - b. What is our best estimate of Jason's GPA at FU?

SUGGESTED READINGS

- Dretzke, B. J., & Heilman, K. A. (1998). *Statistics with Microsoft Excel*. Upper Saddle River, NJ: Prentice Hall.
- Heiman, G. W. (2000). *Basic statistics for the behavioral sciences* (3rd ed.). Boston: Houghton Mifflin.
- Kirk, R. E. (1999). *Statistics: An introduction* (4th ed.). Fort Worth, TX: Harcourt Brace.
- Runyon, R. P., Coleman, K. A., & Pittenger, D. J. (2000). *Fundamentals of behavioral statistics* (9th ed.). New York: McGraw-Hill.
- Stevens, S. S. (1951). *Handbook of experimental psychology*. New York: Wiley.
- Sweet, S. A. (1999). *Data analysis with SPSS*. Needham Heights, MA: Allyn & Bacon.